An Empirical Investigation of Kaizen Event Effectiveness: Outcomes and Critical Success Factors

Jennifer A. Farris

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Industrial and Systems Engineering

Dr. Eileen Van Aken (Chair)
Dr. Kimberly Ellis
Dr. C. Patrick Koelling
Dr. Richard Groesbeck
Dr. Geoffrey Vining
Dr. Toni Doolen

December 18, 2006
Blacksburg, Virginia

Keywords: Kaizen, Kaizen Event, Teams, Lean Manufacturing

An Empirical Investigation of Kaizen Event Effectiveness: Outcomes and Critical Success Factors

Jennifer A. Farris

ABSTRACT

This research presents results from a multi-site field study of 51 Kaizen event teams in six manufacturing organizations. Although Kaizen events have been growing in popularity since the mid 1990s, to date, there has been no systematic empirical research on the determinants of Kaizen event effectiveness. To address this need, a theory-driven model of event effectiveness is developed, drawn from extant Kaizen event practitioner articles and related literature on projects and teams. This model relates Kaizen event outcomes to hypothesized key input factors and hypothesized key process factors. In addition, process factors are hypothesized to partially mediate the relationship between input factors and outcomes. Following sociotechnical systems (STS) theory, both technical and social (human resource) aspects of Kaizen event performance are measured. Relationships between outcomes, process factors and input factors are analyzed through regression, using generalized estimating equations (GEE) to account for potential correlation in residuals within organizations.

The research found a significant positive correlation between the two social system outcomes (attitude toward Kaizen events and employee gains in problem-solving knowledge, skills and attitudes). In addition, the research found significant positive correlations between the social system outcomes and one technical system outcome (team member perceptions of the impact of the Kaizen event on the target work area). However, none of the three technical system outcomes (employee perceptions of event impact, facilitator ratings of event success and actual percentage of team goals achieved) were significantly correlated.

In addition, the research found that each outcome variable had a unique set of input and process predictors. However, management support and goal difficulty were a common predictors of three out of five outcomes. Unexpected findings include negative relationships between functional diversity, team and team leader Kaizen event experience, and action orientation and one or more outcomes. However, many of the findings confirmed recommendations in Kaizen event practitioner articles and the project and team literature. Furthermore, support for the mediation hypothesis was found for most outcome measures. These findings will be useful both for informing Kaizen event design in practicing organizations and for informing future Kaizen event research.

ACKNOWLEDGEMENTS

There are many people who have contributed to this research, both directly and by providing the friendship, encouragement and support that enabled me to keep going and to enjoy my graduate education. First, I would like to thank my advisor, Dr. Van Aken. Thank you for all the ways you have mentored, encouraged and taught me over the years. Thank you for giving me opportunities to work on many interesting research projects and for helping me develop skills related to research, teaching and other aspects of life in academia. Thank you for acting as my professional advocate and for helping me pursue opportunities for research, teaching, funding, conference presentations, publications and jobs. Thank you also for acting as my personal mentor and for helping me protect and balance my time throughout my graduate studies.

To each of my committee members, thank you for your guidance and support in this dissertation and my graduate studies as a whole. To Dr. Ellis, thank you especially for your mentoring and support in my three semesters as your teaching assistant. You taught me a lot about what it means to be a professor and encouraged me to continue my graduate work. To Dr. Koelling, thank you for your wisdom and guidance in the development of my dissertation framework and for your encouragement to focus on the "big picture" in research and life. To Dr. Groesbeck, thank you for your guidance and support in the many lab research projects we worked on together. You helped me develop many different research skills and have been particularly instrumental in the refinement and analysis of the survey scales for this research. To Dr. Vining, thank you for your kind encouragement and for challenging me to think deeply about the complex statistical issues in this research. Your guidance was instrumental in the development of my analysis methods. To Dr. Doolen, thank you for your faith in me and for the opportunity to work on the Kaizen research program from the conceptual stages. I learned a lot from you about the proposal writing process and also about the specific content areas.

To all of the team members and facilitators who participated in this research, thank you for your willingness to share your time and knowledge to expand the Kaizen event body of knowledge. A special thank you to the facilitators at the six participating organizations, who served as data collection coordinators for their organizations. In addition, I am grateful to the National Science Foundation for supporting this research under grant no. DMI-0451512. Also, a special thank you to June Worley who supervised data collection and data entry for the two organizations on the west coast. Thank you for all your work and for your quick and cheerful responses to any questions that I had.

To all the faculty and staff of the Virginia Tech Industrial and Systems Engineering Department, thank you for your help over the years. I particularly want to thank Lovedia Cole for welcoming me to the department and for all her help in understanding and completing the requirements of a graduate education, Kim Ooms for all her help with scheduling and creating documents for meetings, and Nicole Lafon for all her help with payroll and travel reimbursements. To all the Management Systems graduate students, thank you for all your camaraderie, knowledge and encouragement – both professional and personal – a special thank you to James Glenn and Michael Schwandt who had helped me stay focused (and sane) during these last few months of my dissertation work.

To my family – Mike, Karen, Stephen, Amy, Valerie, Jordan, and Norma – thank you for your constant encouragement and for your faith in me throughout this process. I could not and would not have done this without you. Also, thank you to my Virginia Tech family, the Graduate Christian Fellowship. To Patrice Esson, Mary Dean Coleman, Abbie McGhee, Jay McGhee, Amy Albright, Dustin Albright, Shannon Alford, Jarrod Alford and Julia Novak, especially, thank you for your friendship, fellowship and wisdom throughout the years and for helping me have fun and stay focused in graduate school and life. Finally, I am grateful to God who makes all things possible.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

An Empirical Investigation of Kaizen Event Effectiveness: Outcomes and Critical Success Factors

CHAPTER 1: INTRODUCTION

1.1 Research Motivation

A “Kaizen event” is a short-term, team-based improvement project focused on eliminating waste in and increasing the performance of a specific process or product line, through low cost, creativity-based solutions (e.g., Melnyk et al., 1998; Bicheno, 2001). Kaizen events are often associated with the implementation of lean production practices (Vasilash, 1997; Kirby & Greene, 2003) and often employ lean concepts and tools – such as single minute exchange of die (SMED), value stream mapping (VSM), work standardization and 5S (Bodek, 2002; Melnyk et al., 1998; Oakeson, 1997). For more information on lean concepts and tools see Monden, 1983; Womacket al., 1990; and Womack and Jones, 1996b.

From the current Kaizen event body of knowledge, it appears that the intended impact of any given Kaizen event is twofold: first, to substantially improve the performance of the targeted work area, process, or product; and second, to develop the underlying human resource capabilities – the employee knowledge, skills and attitudes (KSAs) – needed to create an organizational culture focused on continuous improvement in the long-term (Sheridan, 1997b; Melnyk et al., 1998; Laraia et al., 1999). A Kaizen event contains both a technical system – i.e., tasks, equipment, and target work area, process, or product – and a social system – i.e., personnel and workforce coordination policies. Thus, a Kaizen event can be studied under the sociotechnical systems (STS) framework (Pasmore & King, 1978). In addition, the improvements intended to be achieved through Kaizen events occur both in the technical system – e.g., improvements in cycle times, WIP, etc. in the target work area – and the social system – e.g., positive changes in employee knowledge, skills and attitudes, etc.

Published accounts of Kaizen event activities, while anecdotal in nature, suggest that Kaizen events can produce rapid and substantial improvement in the technical systems of the work area, processes and products targeted. For instance, one company reported an 885% increased in productivity within one work area (Sheridan, 1997b). Many other organizations have reported significant improvements – often 50% or greater – in key operating measures such as lead-time, floor space, work in process (WIP), throughput/cycle time, productivity, on-time delivery rate, and defect rate (Vasilash, 1993; Redding, 1996; Rusiniak, 1996; Sheridan, 1997b; Oakeson, 1997; Cuscela, 1998; Melnyk et al., 1998; Minton, 1998; LeBlanc, 1999; McNichols et al., 1999; Hasek, 2000; Creswell, 2001;

Butterworth, 2001; Bane, 2002; Bradley & Willett, 2004; Martin, 2004). However, it is also important to note that the published accounts of Kaizen event technical system results appear to come from a biased "sample" – the anecdotal published results of improvements appear to focus on the most successful events.

In addition to their apparent potential for generating rapid and substantial improvement in the technical system performance of the target work areas, processes and products in key business metrics such as throughput, cycle time, lead-time, WIP, etc., it has been suggested that Kaizen events can result in desirable human resource – e.g., social system – outcomes for participating employees. However, in contrast to technical system benefits, specific measures of social system benefits have rarely been reported in the current, anecdotal Kaizen event body of knowledge (Kosandal & Farris, 2004; Miller, 2004). Thus, there is very little evidence, even anecdotally, that Kaizen events are successful in achieving these purported improvements. Nevertheless, some of the favorable social system results claimed (but not demonstrated) are: *enthusiasm for Kaizen event participation* (Tanner & Roncarti, 1994; Wittenberg, 1994; Rusiniak, 1996; Heard, 1997; Taninecz, 1997; McNichols et al., 1999; David, 2000; Hasek, 2000; Bicheno, 2001; Kleinsasser, 2003; Kumar & Harms, 2004); *"buy-in" for change* (Taylor & Ramsey, 1993; Melnyk et al., 1998; Wheatley, 1998; McNichols et al., 1999; Bradley & Willett, 2004); *increased employee empowerment* (Tanner & Roncarti, 1994; Sheridan, 1997b; "Keys to Success," 1997; Melnyk et al., 1998; Minton, 1998; McNichols et al., 1999); *increased employee skills in problem-solving tools and techniques* (McNichols et al., 1999; Watson, 2002); and *development of a culture that supports long-term improvement* (Tanner & Roncarti, 1994; Adams, et al., 1997; "Keys to Success," 1997; Melnyk et al., 1998; Minton, 1998; Foreman & Vargas, 1999; Laraia, et al., 1999; David, 2000; Smith, 2003; Bradley & Willett, 2004; Drickhamer, 2004b; Jusko, 2004; Martin, 2004).

The apparent potential for organizational improvement from the use of Kaizen events is impressive. It appears that many organizations have responded to this potential, as recent evidence suggests that Kaizen events are becoming increasingly popular (Oakeson, 1997; Melnyk et al., 1998; Bodek, 2002; Bane, 2002). However, despite their popularity and apparent potential, there is a dearth of systematic research on Kaizen events. The current Kaizen event body of knowledge is almost entirely practitioner-focused. However, even the few research articles (e.g., Melnyk et al., 1998) focus primarily on describing and defining Kaizen events, rather than building theory to explain Kaizen event effectiveness. The current Kaizen event body of knowledge is focused on anecdotal results from companies that have implemented Kaizen events (e.g., Sheridan, 1997b; Cuscela, 1998) and unproven design recommendations from individuals and organizations that facilitate Kaizen events (e.g., Vitalo, et al., 2003; Mika, 2002; Laraia et al., 1999), rather than on seeking systematic, theory-based understanding of Kaizen events. Thus, there is no systematic, empirical evidence of what sort of Kaizen event designs may be most effective for achieving positive technical system and/or social system outcomes. Although such guidelines would be very useful for practitioners, there are no research-based guidelines for how organizations can compensate for less desirable organizational or work area characteristics through event design. Therefore, there is a strong need for additional, systematic research on Kaizen events. The next section describes how the guiding questions for the research were used to identify ways to address a part of this research need.

1.2 Research Questions

The overall guiding research question for the research was: *What are the technical and social system outcomes of Kaizen events, and what are the factors influencing these outcomes?* Several sub-questions stem from this overall question, including:

- *What is the range of technical system outcomes achieved through Kaizen events, both within a single organization and across organizations?*
- *What is the range of the social system outcomes achieved through Kaizen events?* Do Kaizen events appear to produce the purported positive changes in employee KSAs that are aligned with continuous improvement, as well as with sustaining the Kaizen event improvement program over time?
- *What is the relationship between technical system outcomes and social system outcomes for a given Kaizen event?* Do those events that produce a greater improvement in the technical system also produce a greater improvement in the social system, and vice versa?
- *Finally, what event input and event process factors are most strongly related to event outcomes?* For instance, do the most influential factors differ for technical system outcomes versus social system outcome?

The only way to adequately address these and other research questions is to systematically study Kaizen events in real organizations. Therefore, the research reported here used a multi-site field study to sample Kaizen events in participating organizations to test and refine a working theory of Kaizen event effectiveness.

1.3 Research Purposes and Objectives

The primary purpose of the research is to increase understanding of how Kaizen events operate – that is, to increase understanding of the relationship between event input factors, event process factors and event outcomes, by developing and testing a working theory of Kaizen event effectiveness. This research will fulfill needs of both researchers and organizations by addressing a gap in the current Kaizen event body of knowledge. The overall research purpose led to the identification of the following major research objectives:

- *To describe the technical system outcomes of Kaizen events.*
- *To describe the social system outcomes of Kaizen events.*
- *To identify the relationship between technical and social system outcomes.*
- *To identify the event input factors and event process factors that are related to event outcomes*. The initial stages of the research led to the development of a preliminary theory of Kaizen event effectiveness, based on the published Kaizen event articles as well as other related theory – i.e., project success factor theory, team effectiveness theory, etc. This preliminary theory is graphically depicted in the research model (Figure 1). The preliminary theory was empirically tested using a sample of Kaizen events from organizations participating in the research. Research results and refinements to the theory are discussed in Chapters 3 through 7.
- *To develop guidelines for practitioners managing Kaizen events*. The relationships identified in the research will be translated to actionable guidelines for practitioners seeking to design more effective Kaizen events.
- *To provide tools and methods for organizations and researchers to use to measure event input factors, event process factors, and event outcomes.* The tools and methods used to measure Kaizen event outcomes, process factors and input factors in the research can also be used by organizations to improve their Kaizen event process (McGarrie, 1998; Jha, et al., 1999; Beckett et al., 2000; Bititci, et al., 2000; Gulbro et al., 2000), as well as used by other researchers to study Kaizen events.
- *To provide an empirically tested theory for studying and managing Kaizen events.* The results of this research have been used to refine the working theory of Kaizen event effectiveness. This refinement also provides direction for further research.

1.4 Problem Statement

The research used a multi-site field study to empirically test and refine a preliminary working theory of Kaizen event effectiveness, to increase understanding of the way Kaizen events operate within organizations. This was achieved by measuring technical system outcomes, social system outcomes, event input factors and event process factors for a sample of Kaizen events within participating organizations, and empirically investigating the relationship between these variables for the events within this sample. While Kaizen events appear to be increasingly popular as an organizational improvement mechanism, the currently available guidelines are not based on empirical research, and it is not clear what relationship, if any, exists between the numerous design suggestions and event outcomes. Furthermore, there has been no research to demonstrate whether Kaizen events are successful in producing the favorable changes in employee capabilities that have been claimed in the practitioner literature. Finally, there has been only anecdotal documentation of the technical system results of Kaizen events both within and across organizations. The study of less successful events, in particular, appears to have been neglected to date, as has been documentation of the range of improvements. Empirical investigation of the outcomes achieved through Kaizen events, as well as the relationship between event input factors, event process factors and outcomes will help organizations and researchers better understand, design and manage Kaizen events.

1.5 Sub-Problems and Outputs

The problem statement will be addressed by breaking the problem into the following sub-problems. Each of the sub-problems will result in an output that will further the achievement of the overall research objectives. The following sub-problems are considered in this research:

- *Sub-Problem 1: Identify technical system outcomes, social system outcomes, event input factors and event process factors of interest to the research.* This was achieved through a review of Kaizen event articles, as well as related theory – e.g., sociotechnical systems (STS), industrial and organizational (I/O) psychology, organizational change, project management, and team effectiveness. Outcomes were specified primarily through the review of the published Kaizen event articles and identification of measurement frameworks from STS theory. Outcomes were initially specified as part of a broader Oregon State University (OSU) - Virginia Tech (VT) initiative to study Kaizen events, and the outcomes specification was further refined as a part of this research. To identify potentially relevant input factors and process factors, both a "top down" and "bottom up" approach were used. The "top down" approach was the review of theory in related fields (e.g., project management, team effectiveness and organizational improvement) to identify factors likely to impact Kaizen event outcomes. For instance, Kaizen events can be defined as short-term improvement projects, while cross-functional teams are the human resource structure of Kaizen events. Thus, the project success factor and team effectiveness theory can inform the study of Kaizen events. In addition, a preliminary global research framework developed during pilot research in the broader Oregon State University (OSU) – Virginia Tech (VT) study of Kaizen event effectiveness was consulted as part of the "top down" approach. The "bottom up" approach included the review of the largely anecdotal, Kaizen event articles to identify factors mentioned in these sources as potentially impacting outcomes.

- *Sub-Problem 2: Develop a working theory of Kaizen event effectiveness. That is, identify a subset of relevant event input factors, event process factors, and outcomes, and specify the relationship between these constructs.* The preliminary working theory was developed using the results of the previous sub-problem and critical thinking by the researcher. The research model (Figure 1) describes the preliminary theory of Kaizen event effectiveness — i.e., the hypothesized relationships between event input factors, event process factors, and outcomes – this theory was tested and refined in the research (see Sub-Problem 3).
- *Sub-Problem 3: Test the hypothesized relationships between event input factors, event process factors, and outcomes.* This sub-problem involved collecting and analyzing data from Kaizen events. This included identifying ways to measure the event input factors, event process factors and outcomes of interest, ways to select events for study, and ways to analyze the resulting data to test hypotheses..
- *Sub-Problem 4: Refine the working theory of Kaizen event effectiveness.* The relationships identified in the previous sub-problem were interpreted to refine the working theory of Kaizen event effectiveness.
- *Sub-Problem 5: Develop guidelines for organizations and researchers, based on the outcomes of this research.* This sub-problem involved interpreting the results of the data analysis (Sub-Problem 3) and the revised research model (Sub-Problem 4) to develop guidelines for organizations seeking to design and manage more effective Kaizen events.

The following are the outputs of the sub-problems.

- *Output of Sub-Problem 1: Documentation of the technical system outcomes, social system outcomes, event input factors and event process factors of interest to the research.* Section 1.6 and Chapter 2 describe the technical system and social system outcomes, and event input and process factors of interest.
- *Output of Sub-Problem 2: A working theory of Kaizen event effectiveness, depicted visually in the preliminary research model (Figure 1).* The research model describes the hypothesized relationships between the event input factors, event process factors and outcomes that were studied in the research.
- *Outputs of Sub-Problem 3: Specification of the relationships between event input factors, event process factors, and outcomes.* This includes reporting the strength of the observed relationship between event input factors, event process factors and outcomes – i.e., the results of the hypothesis tests. An auxiliary output of this sub-problem is the identification, development and refinement of *tools for measuring event input factors, event process factors, and technical system and social system outcomes*, which can be used by organizations conducting Kaizen events and researchers studying Kaizen events or similar change mechanisms.
- *Output of Sub-Problem 4: A refined, working theory of Kaizen event effectiveness.* A revised visual depiction – e.g., research model – is presented in Chapter 6 (see Figure 8).
- *Output of Sub-Problem 5: Guidelines for organizations for designing and managing more effective Kaizen events.* Guidelines include recommendations about which variables are the key input factors and event process factors that should be managed to achieve favorable Kaizen event outcomes, as well as how interrelationships between input factors and event process factors can be managed.

1.6 Research Model and Definitions

In order to further the research purpose of developing and testing a working theory of Kaizen event effectiveness, the following preliminary research model (Figure 1) was developed. The research model graphically depicts the preliminary working theory of Kaizen event effectiveness – i.e., the hypothesized relationships between study variables. The model began as a conceptual research model, defining broad categories of variables. This initial model was refined into an operational research model as the literature review (Chapter 2) and research design (Chapter 3) were completed.

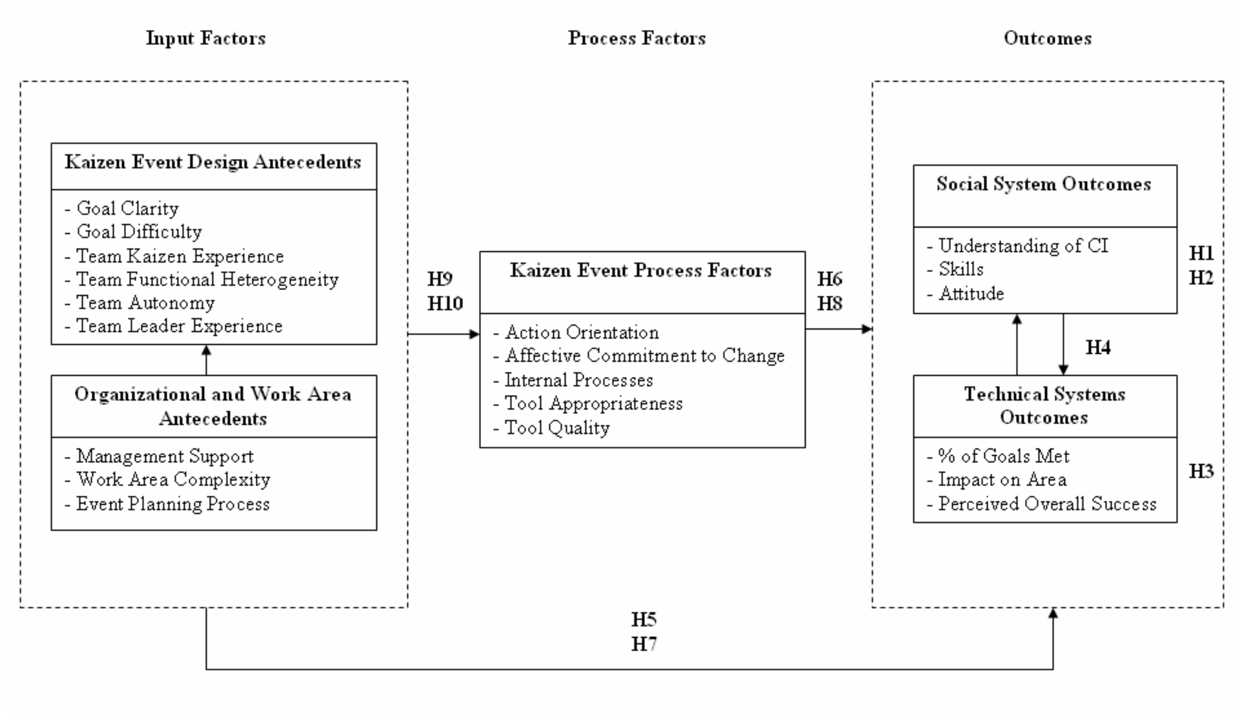

Figure 1. Preliminary Operational Research Model

Kaizen Event Design Antecedents represent input factors that describe the design of a given Kaizen event and can be directly manipulated by the individual designing the Kaizen event. In many organizations, there is an individual dedicated full-time to the role of designing and implementing Kaizen events. This individual is often called the "Kaizen event coordinator." The variables in this category studied in this research are:

- *Goal Clarity* —this variable describes team member perceptions of the extent to which the Kaizen event team's improvement goals have been clearly defined.
- *Goal Difficulty*– this describes team member perceptions of the difficulty of the improvement goals set for the Kaizen event team.
- *Team Kaizen Experience* – this variable describes the average Kaizen event experience within the Kaizen event team
- *Team Functional Heterogeneity* – this variable describes the diversity of functional expertise within the Kaizen event team
- *Team Autonomy* – this variable describes the amount of control over event activities that is given to Kaizen team members.

- *Team Leader Experience* – this variable describes the total number of Kaizen events that the team leader has led or co-led.

Organizational and Work Area Antecedents include input factors that describe the context of the event, and were hypothesized to influence Kaizen event design antecedents, as well as to directly influence the Kaizen event process and outcomes. The variables in this category studied in this research are:

- *Management Support* – this variable describes the support that senior leadership provided to the team, including materials and supplies, equipment and assistance from organizational members – the facilitator, senior management and others.
- *Event Planning Process* – this variable describes the total person-hours invested in planning the event
- *Work Area Routineness* – this variable measures the general complexity of the target system, based on the level of stability of the product mix and degree of routineness of product flow.

Kaizen Event Process Factors describe the attitudes and activities of the team during the event. The variables in this category studied in the research are:

- *Action Orientation* – this describes the activities of the event team in terms of the relative percentage of team activities spent in hands on improvement activities – e.g., in the target work area or process – versus brainstorming and discussing solutions in offline team meetings – e.g., in meeting rooms.
- *Affective Commitment to Change* – this describes team member perceptions of the need for the specific changes targeted by the Kaizen event.
- *Tool Appropriateness* -- this variable describes the average facilitator rating for the appropriateness of the problem solving tools selected by the team.
- *Tool Quality* -- this variable describes the average facilitator rating for the quality of the problem solving tools selected by the team.
- *Internal Processes* – this variable describes team member ratings of the internal harmony and coordination of their team.

Technical System Outcomes describe the technical improvements to the target system. The variables in this category studied in this research are:

- *% of Goals Met* – the aggregate percentage of main improvement goals met.

- *Overall Perceived Success* – this variable describes stakeholder perceptions of the overall success of the Kaizen event.
- *Impact on Area* – this variable describes team member perceptions of the impact of the Kaizen event on the target system.

Social System Outcomes describe changes to team member KSAs that are aligned with continuous improvement. The variables in this category studied in this research are:

- *Understanding of Continuous Improvement (CI)* – this variable describes the degree to which team members' knowledge of the philosophy of continuous improvement was improved by the Kaizen event
- *Attitude* – this variable describes the degree to which team members' liking for Kaizen event activities was increased by the event.
- *Skills* – this variable describes that degree to which team members report gaining skills from the Kaizen event activities.

1.7 Research Hypotheses

To meet the objectives of the research, the following hypotheses were investigated. These hypotheses were developed based on the research model (Figure 1), which describes the hypothesized relationships between study variables:

H1. Social system outcome variables will be significantly correlated at the team level.

H2. Social system outcomes will occur primarily at the team level, rather than individual level, indicated by significant intraclass correlation for social system outcome variables.

H3. Technical system outcome variables will be significantly correlated at the team level.

H4. Social system outcomes will be significantly correlated with technical system outcomes at the team level.

H5. Event input factors will be positively related to social system outcomes at the team level.

H6. Event process factors will be positively related to social system outcomes at the team level.

H7. Event input factors will be positively related to technical system outcomes at the team level.

H8. Event process factors will be positively related to technical system outcomes at the team level.

H9. Event process factors will partially mediate the relationship of event input factors and social system outcomes at the team level.

H10. Event process factors will partially mediate the relationship of event input factors and technical system outcomes at the team level.

1.8 Overview of Research Design, Premises, and Delimitations

The research design is a multi-site field study that sampled Kaizen events within participating organizations in order to test a working theory of Kaizen event effectiveness. This empirical investigation allowed the analysis of the interrelationships between technical system outcomes, social system outcomes, event input factors and event process factors in a way that other study methods would not. For instance, organization-wide survey methods could be used to study the social system outcomes of Kaizen events. However, this would not provide a clear link between social system outcomes and specific Kaizen events. In addition, organization-wide sampling of social system outcomes combined with post-hoc sampling of technical system outcomes – e.g., for past events – also would not allow the accurate analysis of the relationship between technical and social system outcomes, because the measurement of technical and social system outcomes would not be concurrent. In addition, while the accurate, post-hoc measurement of certain event input factors and event process factors may be possible – e.g., team functional heterogeneity – the accurate, post-hoc measurement of other variables would not be possible due to recall difficulties and history effects. A field study that samples Kaizen events as they occur, therefore, is an appropriate method that allows accurate, simultaneous measurement of the technical system outcomes, social system outcomes, event input factors and event process factors from specific Kaizen events.

The following premises guided the design of the research:

- Studying Kaizen event teams in the field is preferable to studying Kaizen events in an artificial environment – e.g., lab experiments – in terms of generalizability (Cohen & Bailey, 1997).

The following limitations of scope apply to this research:

- This research did not attempt to study all potential technical system and social system outcomes. While the potential technical system outcome measures can be fairly narrowly defined – e.g., initial outcomes, outcomes over time, etc. – the variety of social system outcomes that could be impacted by Kaizen events is large. The Kaizen event body of knowledge and related bodies of knowledge were reviewed to identify a parsimonious set of social system outcome measures that appear to be particularly relevant to the study of Kaizen events. These measures were identified as the social system dependent variables for this research. Other measures – e.g., employee satisfaction, etc. – could be the subject of future research.

- This research did not attempt to study all event input factors and event process factors that could potentially impact Kaizen event outcomes. Instead, the factors chosen for this research were selected from the Kaizen event body of knowledge and related theory as dominant, recurring factors indicated by multiple sources as likely determinants of event outcomes. One purpose of this research was to evaluate the fit of the preliminary research model and to suggest whether additional factors not studied in the current research should be investigated in future research. In addition, there are other methods of identifying potentially important event input and event process factors that were not applied in this research. For instance, the research could interview Kaizen event facilitators/coordinators and other organizational personnel to identify what they believe to be the most important factors that influence Kaizen event outcomes. This method is known as critical success factors (CSF) interviews (Rockart, 1979). This method appears to be particularly useful when very little has been published on a given subject (Nicolini, 2002), and when it is difficult to empirically sample the entities in question due to time and resource constraints. Due to the availability of published, albeit anecdotal, accounts of purported relevant factors from the practitioner literature and related branches of theory, as well as the opportunity for sampling multiple Kaizen events due to their short duration, the CSF interview method was not used in the research. However, this could be the subject of future research. For instance, a follow-up study could be conducted to confirm whether the factors identified in this study as most strongly related to event outcomes are aligned with practitioner perceptions. In addition, the Kaizen Event Program Interviews which were conducted as an auxiliary step in this research provide some, albeit sometimes indirect, insight into the factors that Kaizen event coordinators at the participating organization believe are important.
- This research only studied the initial impact of Kaizen events. Research on the longitudinal sustainability of both technical and social system outcomes is needed, but in-depth treatment of this subject is outside the scope of this research. However, the longitudinal sustainability of event outcomes, and the relationship of sustainability to Kaizen event follow-up factors and organizational and work area antecedents, is included in the broader OSU - VT initiative to understand Kaizen events.
- This research had a fairly small sample size at the organizational level (i.e., six organizations). In addition, the organizations are all manufacturing organizations of some type, although they apply Kaizen events to non-

manufacturing processes. The organizations are also limited geographically to within a day's drive of either VT or OSU. This allowed the possibility of site visits, while controlling costs. These site visits were used to train organizational personnel in the data collection methods and to better understand the context of the data collected, not for the direct collection of study data. Therefore, it is possible that study results will not be generalizable to organizations with markedly different characteristics to the organizations included in this research – e.g., organizational context and/or climate; industry sector; region, etc. However, at an event and individual level, the diversity of the events studied increases the support for generalizability of the research findings, even with the relatively small sample size at the organizational level. A variety of events – which involved a variety of goals, processes, participants and target work areas – were studied. For instance, the study of Kaizen events in non-manufacturing processes within the participating organizations increases the support for generalization to non-manufacturing organizations. In addition, contextual data about participating organizations will be collected to identify unique characteristics likely to impact generalizability.

1.9 Contributions of this Research

The research contributes to the general body of knowledge concerning Kaizen events, to organizational practice and to the discipline of industrial engineering. This research contributes to the body of knowledge by systematically and empirically documenting the outcomes of Kaizen events, and by identifying the event input factors and event process factors that are most strongly related to event outcomes. While many examples of Kaizen event benefits have been cited in the practitioner literature, there has been no empirical research to verify the types of technical and social system outcomes typically achieved through Kaizen events. In particular, social system outcomes appear to be rarely measured. In addition, although numerous design suggestions for managing Kaizen events exist, these practitioner guidelines have not been systematically tested. Finally, no explicit link has been made between Kaizen events and related organizational theory – e.g., team effectiveness, project management, etc.. This research contributes an empirically-tested theory for studying and managing Kaizen events, which describes the types of outcomes achieved through Kaizen events, the relationship between event input and process factors and event outcomes, and the relationship of Kaizen events to related organizational structures – i.e., projects and teams. The results of the research will be used to create theory-based, qualitative guidelines for managing Kaizen event effectiveness. Finally, the data collection tools and techniques develop for this research provide tools and methods for organizations and researchers to use to measure event input factors, event process factors, and event outcomes.

This research contribute to organizational practice by increasing understanding of the factors that effect Kaizen event outcomes. Thus, the research will help organizations design and implement more effective Kaizen events. In addition, an understanding of the social system outcomes can help the organization in its longer-term development of its human resources. The relationships identified by the research will be summarized as recommended guidelines for practitioners conducting events under different scenarios.

This research clearly contributes to the discipline of industrial engineering. The Institute of Industrial Engineers (IIE) has adopted the following definition of industrial engineering:

> "Industrial Engineering is concerned with the design, improvement, and installation of integrated systems of people, material, information, equipment, and energy. It draws upon specialized knowledge and skills in the mathematical, physical, and social sciences together with the principles and methods of engineering analysis and design to specify, predict, and evaluate the results to be obtained from such systems." (Institute of Industrial Engineers, 2006).

As a discipline, industrial engineering is aimed at the design and improvement of integrated systems. A Kaizen event is an integrated system of people, materials, information, equipment and energy. By identifying Kaizen event input factors and event process factors that contribute to successful Kaizen event outcomes, and by developing, testing and refining a working theory of Kaizen event effectiveness, the research will enable the improvement of Kaizen events (as integrated systems), by providing suggestions for design, as well as implementation.

Furthermore, Kaizen events, as temporary integrated systems, are designed to improve more permanent integrated systems. Thus, any findings that improve the effectiveness of Kaizen events as a system will also lead to the improvement of other systems within the organization – e.g., specific production systems, employee development, etc. In addition, Kaizen events can also be used as a vehicle to design and/or install organizational systems.

Methodologically, the study also clearly falls within the discipline of industrial engineering. Not only does the research seek to provide findings to enable the improvement of integrated systems, the research required the application of specialized knowledge in the mathematical, physical and social sciences. Through the use of these industrial engineering methods, the research seeks to define and evaluate the results obtained from integrated Kaizen event systems, and to support the prediction of results from future events by identifying the relationships between event input factors, event process factors and event outcomes.

The remainder of this document is organized as follows. Chapter 2 provides a review of the key articles and theories used to develop the initial research model and to specify study hypotheses –i.e., to develop the working theory of Kaizen event effectiveness. Chapter 3 describes the methods used to collect study data and to prepare the data for hypothesis testing. Chapter 4 presents the analysis methods used to test the study hypotheses and the results of the hypothesis tests. Chapter 5 provides a discussion of the study results and implications for researchers and practicing organizations. Chapter 6 summarizes the findings of this research and describes areas of future research.

CHAPTER 2: LITERATURE REVIEW

The following chapter reviews the bodies of knowledge that were examined as a foundation for understanding the relationships between event input factors, event process factors and event outcomes in order to build the working theory of Kaizen event effectiveness. Unlike many dissertation literature reviews, this review is organized by research model component, rather than strictly by body of knowledge. First, a review of the body of knowledge related to Kaizen event outcomes is presented. Second, a review of the body of knowledge related to input factors and process factors is presented. Next, this chapter describes how a preliminary global research framework developed during pilot research in the broader Oregon State University (OSU) – Virginia Tech (VT) study of Kaizen event effectiveness also informed this study. Finally, the specification of the initial research model is described.

Both the Kaizen event process and Kaizen event outcomes can be conceptualized under the STS framework (Trist & Bamforth, 1951; Emery & Trist, 1960; Trist et al., 1963; Katz & Kahn, 1966; DeGreene, 1973; Pasmore & King, 1978). For instance, to achieve improvement in the target system, the Kaizen event process uses a social system – e.g., a relatively autonomous cross-functional team – as well as a technical system -- e.g., a suite of performance improvement tools and procedures. In addition, Kaizen events are purported to produce desirable change in both a technical system – e.g., the target work area, process or product – and a social system – e.g., the Kaizen event team and employees who work in the target work area or process), which coincides with the STS principle of joint optimization. Thus, STS theory served as an overall guiding framework in the identification of relevant types of input factors, process factors and outcomes.

2.1 Review of the Literature Related to Kaizen Event Outcomes

2.1.1 Introduction to Kaizen Events

Although there is a lack of research literature on Kaizen events, the practitioner literature describes the overall concept of and intended outcomes of Kaizen events. In addition, an older body of literature describes the concept of *kaizen* – that is, is continuous, incremental improvement of all aspects of an organization (Imai, 1986).

A Kaizen event is a "short-term project focused on a specific process or set of activities, such as the work flow within a specific work center" (Melnyk et al., 1998, p. 70). A complementary definition describes a Kaizen event as "a focused improvement event during which a cross-functional team of operators, engineers, etc. spends several days analyzing and implementing improvements to a specific work area" (Kirby & Greene, 2003, p. 2). Another definition describes a Kaizen event as "rapid, repeated time-compressed changes for the better in bite-sized chunks of the business" (Heard, 1997, p. 522-523). Other terms used in the evolving literature to represent this improvement mechanism include: "kaizen blitz" (Cuscela, 1998); "kaikaku" (Bicheno, 2001; Womack & Jones, 1996a); "rapid kaizen" (Melnyk et al., 1998); "breakthrough kaizen" (Womack & Jones, 1996b); "Gemba kaizen" (Wittenberg, 1994; Sheridan, 1997a; Sabatini, 2000; Mika, 2002); "kaizen workshops "(Sheridan, 1997b), "short cycle kaizen" (Heard, 1997; Heard, 1998), and "just do it (JDIT) – kaizen" (Taylor & Ramsey, 1993).

The growth in the practitioner literature seems to indicate that Kaizen events began gaining in popularity in the mid- to late- 1990s. However, one source indicates that Toyota used rapid-change projects similar to Kaizen events with suppliers as far back as the 1970s (Sheridan, 1997b). Most Kaizen events are conducted in the format of a three to five day work session, although some Kaizen events are shorter, one or two-day work sessions (Sheridan, 1997b; Cuscela, 1998; Minton, 1998). Typical activities include training, documenting current processes, identifying opportunities for improvement, implementing and evaluating changes, presenting results to management, and developing an action plan for future improvements (Melnyk et al., 1998). Generally, Kaizen events use a cross-functional team of employees solely dedicated to the event for its duration (Minton, 1998; Bicheno, 2001). Events generally occur during business hours and employees appear to receive compensation for their participation as part of their normal salaries or wage, and perhaps even as overtime (when events exceed normal business hours). Although often associated with the implementation of lean production (Sheridan, 1997b; Vasilash, 1997; Melnyk et al., 1998; Sheridan, 2000a; Kirby & Greene, 2003; Bradley & Willett, 2004), Kaizen events can be used to achieve improvement apart from a formal lean production program (Kleinsasser, 2003; Bane, 2002).

The scope of the changes sought during Kaizen events is limited, focused on all or part of a specific process or product line, rather than broad, organization-wide policies or technology changes (Laraia et al., 1999). Improvement goals are set in advance by organizational leadership (Melnyk et al., 1998). In addition to the limitations in breadth, another typical limitation to Kaizen event activities is the limitation on capital investment. A focus of Kaizen events is on making improvements to existing processes or products without (or at least before) investment in new technology. Thus, Kaizen events generally have little or no budget for capital equipment (Sheridan, 1997b; Bicheno, 2001). Instead, the focus is on using human knowledge and creativity, through the application of a systematic problem solving methodology and structured process tools (Bicheno, 2001). Thus Kaizen events stand in stark contrast to other improvement approaches – e.g., business process reengineering – which focus on applying new technology, often requiring significant investment (Harrington, 1998).

Kaizen events typically use cross-functional teams composed of employees from the target area as well as support personnel, such as engineering, management or maintenance. In some cases, customers or suppliers may also be involved directly on a Kaizen event team. Using a cross-functional team increases the variety of knowledge and perspectives that can be applied to the team's problem-solving tasks, and would also appear to build support and ownership for the event-based changes for the Kaizen event team members from the target area (McNichols et al., 1999). Several practitioner sources suggest that about half of the Kaizen event team members should be employees from the target work area (Minton, 1998; Mika, 2002). In addition to generating support for the changes made during the event, involving target area employees on the Kaizen event team could introduce employees to tools and techniques of lean process improvement they would need to effectively sustain the event-based changes, as well as to generate future improvements (Vasilash, 1993; Womack & Jones, 1996b, Bradley & Willett, 2004). Thus, a Kaizen event can be viewed as an investment in employee training (Rusiniak, 1996; Adams et al., 1997; Vasilash, 1997; McNichols et al., 1999). Another apparently unique feature of Kaizen events as an organizational improvement mechanism is their strong action orientation. While many "traditional" improvement projects end with a set of recommended changes presented to senior management, Kaizen teams are generally given the authority to implement solutions as they are developed, without direct approval of each change from management (Oakeson, 1997; Sheridan, 1997b; Minton, 1998; Laraia et al., 1999; LeBlanc, 1999; Bicheno, 2001). In many organizations, at the end of the event, Kaizen event teams are expected to present management with a changed work area, where the team's solution has already been implemented, measured, and is being standardized. Depending on the goals of the event, Kaizen teams are often also encouraged to achieve more than one cycle of solution refinement (Bradley & Willett, 2004; Bicheno, 2001; Melnyk et al., 1998).

While there appears to be no reason why Kaizen events cannot be used as single, isolated improvement mechanisms, many organizations that conduct Kaizen events conduct them very frequently – dozens, or even hundreds, of events per year (Vasilash, 1997; Sheridan, 1997b; LeBlanc, 1999). The practitioner literature also suggests that Kaizen events are intended to be used at multiple points in time, as an ongoing program of continuous organizational improvement (Laraia et al., 1999, Melnyk et al., 1998; LeBlanc, 1999). Multiple Kaizen events can be used to continuously improve a given process or product (Melnyk et al., 1998), or to create improvement in many different areas of an organization. Thus, while taking a different implementation form, Kaizen events are linked to the incremental improvement approach known as "kaizen."

2.1.2 "Kaizen Event" versus "Kaizen"

Although one source indicates that the first Kaizen event-like activities originated with Toyota (Sheridan, 1997b), a recent study on Kaizen in Japan describes Kaizen events as a "more recent Western development" (Brunet & New, 2003, p. 1428). Thus Brunet and New's (2003) study of kaizen in Japanese corporations clearly distinguishes "Kaizen events" from the older concept of "kaizen." Although apparently originating in the U.S. in the 1890s (Schroeder & Robinson, 1991), the concept of continuous improvement, or "kaizen," is most often recognized as one of the key principles of Japanese manufacturing and, in fact, appears to have been practiced primarily in Japan from the 1950s-1970s, before being reintroduced to the U.S. in the 1980s (Jha et al., 1996). Imai (1986) reintroduced the concept of kaizen into popular, management literature (Sheridan, 1997a), defining kaizen as the principle of continually and incrementally improving all aspects of an organization through the extensive involvement of employees at all organizational levels. At the time, kaizen appeared to be a uniquely Japanese concept (Martin, 2004), and Imai cited it as the key ingredient in Japan's manufacturing success. Since Imai's first definition, related interpretations of kaizen have been proposed. The term, literally translated, means "good change," combining the Japanese words "kai," meaning change, and "zen," meaning good (iSixSigma LLC, 2004). Another popular definition of the concept of kaizen is "to take apart and put back together in a better way" (Minton, 1998, p. 19).

A "Kaizen event" is related to the concept of kaizen in several ways. First, both concepts include using process improvement tools and techniques – often, the same tools and techniques that are associated with lean manufacturing – to make desired improvements. Second, both concepts include the aim of ultimately producing an organizational culture focused on ongoing improvement (Imai, 1986; Laraia et al., 1999; Melnyk et al., 1998; Sheridan, 1997b). Third, both concepts include the idea of empowering employees to make changes, by providing both opportunity to improve work systems, as well as training in the tools and techniques needed to make improvements. Finally, both concepts emphasize making relatively incremental changes to improve performance. For instance, Kaizen events have a relatively narrow focus – focused on improving a specific work area, process or product, rather than making radical change to broader, organizational systems. In addition, similarly to kaizen, Kaizen events focus on low-cost changes, rather than changes requiring significant capital investment (Sheridan, 1997a). Kaizen events can also be used multiple times in a given work area to create cycles of improvement in the work area (Melnyk et al., 1998). The incremental nature of kaizen is inherent in the Imai (1986) definition.

However, as has been previously mentioned, Kaizen events are distinct from kaizen. The concept of kaizen is a broader concept related to an organizational culture that is supportive of continuous improvement. As a system, kaizen has often been implemented through quality circles, problem-solving or continuous process improvement (CPI) teams, employee suggestion programs, and other ongoing policies that enable employees to participate in improving their daily work. There is some evidence that some organizations have realized benefits through the implementation of these types of continuous improvement programs (McGarrie, 1998; Jha et al., 1999; Hammersley & Pinnington, 1999; Chow-Chua & Gho, 2000), although many such programs have failed (Sterman et al., 1997; Easton & Jarrell, 1998; Keating et al., 1999; Repenning & Sterman, 2002). Quality circles, problem solving and CPI teams, and employee suggestion programs are all longer-term initiatives than the typical Kaizen event. For instance, CPI teams generally meet for a few hours at a time over weeks or months (Mohr & Mohr, 1983). Employee suggestion programs and other policies are even longer-term. Thus, Kaizen events are clearly distinct from the usual ways in which kaizen is implemented as a system. However, it appears that Kaizen events could be a vehicle to implement the concept of kaizen within an organization (LeBlanc, 1999; Kumar & Harms, 2004). Kaizen events could be a component in an organization's kaizen system, used either with more "traditional" systems such as CPI and employee suggestion programs, or by itself. As has been previously mentioned, Kaizen events support key concepts related to kaizen – including, enabling employees to make changes to their work areas, developing an organizational culture focused on ongoing improvement, and achieving improvement incrementally.

In addition to being aligned with the traditional concept of kaizen, Kaizen events appear to offer at least two noticeable potential benefits over the way kaizen is often implemented in organizations. First, Kaizen events often generate immediately perceivable performance improvements. This immediate return on investment could provide the short-term "wins" many in organizational change literature sources cite as necessary to creating employee buy-in (e.g., commitment) to a given improvement program over the longer-term (Keating et al., 1999; Kotter, 1995). Thus, it is possible that organizations that use Kaizen events with, or instead of, longer-cycle "traditional" CPI programs with similar objectives, may ultimately be more successful in sustaining their ability to produce improvements.

Second, because Kaizen event teams typically have authority to implement changes during the event without direct approval from senior management (Oakeson, 1997; Sheridan, 1997b; Minton, 1998; Laraia et al., 1999;

LeBlanc, 1999; Bicheno, 2001), they often have a high degree of autonomy or "substantative participation" compared with CPI teams (Cohen & Bailey, 1997). In contrast, "traditional" CPI teams, at least as implemented in Western organizations, often have no authority to implement – they merely recommend changes to senior management (Mohr & Mohr, 1983; Cohen & Bailey, 1997; Laraia et al., 1999). This is a form of "consultative participation" (Cohen & Bailey, 1997). This distinction between Kaizen event teams and "traditional" CPI teams is important, since team effectiveness research has shown that substantative participation is related to both positive technical performance outcomes (Campion et al., 1993; Cohen & Ledford, 1994; Gupta et al., 1994; Cohen et al. 1996) and positive social system outcomes – e.g., employee satisfaction (Cordery et al., 1991; Pearson, 1992; Weisman et al., 1993; Cohen & Ledford, 1994; Seers et al., 1995; Cohen et al., 1996) and commitment (Cordery et al., 1991; Cohen et al., 1996). Meanwhile, most studies of consultative participation have shown no relationship between consultative participation and either technical or social outcomes (Steel et al., 1990), or a negative relationship between consultative participation and these outcomes (Marks et al., 1986; Adam, 1991). One study (Griffin, 1988) did show initial gains in satisfaction for quality circle members, but these gains decreased substantially after 18 months and disappeared completely after three years. Another study, which directly compared consultative and substantive participation for teams in two industries (telecommunications and apparel) found that substantative participation was a strong predictor of both technical and social system outcomes – e.g., satisfaction, organizational commitment, and workers' perceptions of quality – while consultative participation was only a weak predictor of organizational commitment for one of the two industries (apparel) (Batt & Appelbaum, 1995). Thus, since they rely on substantative participation, rather than consultative participation, Kaizen events have the potential to produce more favorable technical system outcomes and social system outcomes than "traditional" CPI activities with similar focus.

2.1.3 Technical System Outcomes

An examination of the largely practitioner-focused, Kaizen event literature suggests that the process metrics most commonly targeted for improvement include: *productivity*, *work-in-process* (*WIP*), *floor space*, *throughput*, *lead-time, set-up time*, *part travel time*, *percent on-time delivery*, *defect rate*, *throughput and product design measures* (*price, product line diversity, etc.*) (Kosandal & Farris, 2004). Reported improvements in these metrics vary from moderate improvement – e.g., 25-50% – to significant improvement – e.g., 75-100% – to orders of magnitude improvement (Sheridan, 1997b; Cuscela, 1998). However, due to the lack of systematic, research on Kaizen events, the range of "typical" technical system outcomes is not well understood.

2.1.4 Social System Outcomes

The Kaizen event literature claims that several important social outcomes may be achieved through Kaizen events. These purported outcomes include:

- Enthusiasm for Kaizen event participation (Tanner & Roncarti, 1994; Wittenberg, 1994; Rusiniak, 1996; Heard, 1997; McNichols et al., 1999; David, 2000; Kumar & Harms, 2004) – signified in one case by a waiting list of volunteers (Taninecz, 1997). Employees appear to like Kaizen events (Hasek, 2000) – to find them fun (Bicheno, 2001), and to enjoy giving input to the process improvement (Kleinsasser, 2003).
- Support for the Kaizen event program (Redding, 1996) – e.g., "buy-in" to the effectiveness of the Kaizen event process (Heard, 1997; Melnyk et al., 1998; Watson, 2002).
- Employee "buy-in" to the changes made during the Kaizen event, due to the participation of target area employees on the Kaizen event team (Taylor & Ramsey, 1993; Melnyk et al., 1998; Wheatley, 1998; McNichols et al., 1999; Bradley & Willett, 2004).
- Creation of a belief that change is possible (Tanner & Roncarti, 1994; Butterworth, 2001).
- Increased employee empowerment (Tanner & Roncarti, 1994; Sheridan, 1997b; "Keys to Success," 1997; Melnyk et al., 1998; Minton, 1998; McNichols et al., 1999).
- Improved employee attitudes toward work (Minton, 1998).
- Increased cross-functional cooperation, due to the cross-functional nature of Kaizen event teams (Tanner & Roncarti, 1994; McNichols et al., 1999; Mika, 2002; Bradley & Willett, 2004).
- Support for creating a learning organization (Melnyk et al., 1998), by giving employees tools they can apply to improve their daily work activities (McNichols et al., 1999; Watson, 2002).
- Support for lean manufacturing (Bradley & Willett, 2004).
- Development of a culture that supports long-term improvement (Tanner & Roncarti, 1994; Adams, et al., 1997; "Keys to Success," 1997; Melnyk et al., 1998; Minton, 1998; Foreman & Vargas, 1999; Laraia, et al., 1999; David, 2000; Smith, 2003; Bradley & Willett, 2004; Drickhamer, 2004b; Jusko, 2004; Martin, 2004).
- Creation of a "hands on," "do-it-now" sense of urgency for change/improvement in the entire facility (Tanner & Roncarti, 1994)

- Employee pride in accomplishments made during the Kaizen event (Tanner & Roncarti, 1994)

The first two benefits may be especially important since Keating, Oliva, Repenning, Rockart & Sterman (1999) cite lack of employee commitment to the program as one of the reasons many improvement programs, even those that are initially successful, ultimately fail. In addition, many of the benefits are aligned with sustaining the changes made during a given Kaizen event, as well as supporting continuous improvement throughout the organization. For instance, employee training and skill development has been cited as a critical factor in generating and sustaining organizational improvement (McGarrie, 1998; Gulbro et al., 2000; Jha et al., 1999; Beckett et al., 2000).

2.2 Review of the Literature Related to Input Factors and Process Factors

2.2.1 Project Success Factor Theory

Since its inception as a discipline, a major focus of the project management field has been identifying the critical input factors contributing to project success – e.g., "critical success factors." Although research on critical success factors has been ongoing since the 1960s (Belassi & Tukel, 1996), there is still a lack of agreement on both the critical success factors and the definition of "project success" (Shenhar et al., 2002). Many models of the relationship between critical success factors and project outcomes have been proposed. Project management theory can contribute to the study of Kaizen events, since a Kaizen event can be described a short-term improvement project. For instance, Bane (2002) describes a Kaizen team as a "dedicated project team," while Melnyk et al. (1998) and Bradley & Willett (2004) both describe Kaizen events as "improvement projects."

A Kaizen event also conforms to the project management discipline's definitions of a project. For instance, the Project Management Body of Knowledge (PMBOK®) defines a project as: "a temporary endeavor undertaken to create a unique product or service" (PMBOK® Guide, 2000, p. 4). A Kaizen event is both: 1) *temporary* -- with a clearly defined, limited time frame; and 2) *unique* – A Kaizen event is a one time, unique endeavor. It seems unlikely that an organization would ever repeat a Kaizen event in exactly the same way in a given work area . Even events with similar focus that are "repeated" within the same target work area, process or product would be expected to have different objectives – i.e., specific targets for cycle time, setup time, etc. – organizational and work area contexts, team compositions and team activities.

In addition to the PMBOK® definition, Lewis (2000) defines a project as "a one-time, multi-task job that has clearly defined starting and ending dates, a specific scope of work to be performed, a budget, and a specified level of performance to be achieved" (p. 4). Again, a Kaizen event clearly falls under this definition. It has:

- *Clearly defined starting and ending dates* – The timeframe of a given Kaizen event is clearly defined prior to the start of the event. Generally the length of the Kaizen event is one week or shorter (Vasilash, 1993; Adams et al., 1997; Oakeson, 1997; Patton, 1997; Sheridan, 1997b; Vasilash, 1997; Cuscela, 1998; Melnyk et al., 1998; LeBlanc, 1999; McNichols et al., 1999; Bicheno, 2001; Watson, 2002; Smith, 2003; Bradley & Willett, 2004; Drickhamer, 2004b; Martin, 2004).
- *Clearly defined scope* – The objectives and boundaries of a given Kaizen event are also clearly defined prior to the start of the event (Adams et al., 1997; Sheridan, 1997b; Melnyk et al., 1998; Minton, 1998; Bicheno, 2001; Drickhamer, 2004b; Martin, 2004). In addition, Kaizen event teams are instructed not to go beyond the boundaries of their event. Instead, they are instructed to use a "Kaizen newspaper" to note potential improvements that are beyond the scope of their event, for implementation in future activities (Melnyk et al., 1998; McNichols et al., 1999; "Winning with Kaizen," 2002; Bradley & Willett, 2004; Martin, 2004).
- *Clearly defined budget* – Event budgets, although generally small (Vasilash, 1993; Adams et al., 1997; Sheridan, 1997b; Cuscela, 1998; Melnyk et al., 1998; Bicheno, 2001; Martin, 2004; Purdum, 2004), are also defined in advance of the given Kaizen event.
- *Clearly defined performance levels* – Clearly specified, measurable goals is one of the key characteristics of Kaizen events (Martin, 2004; Bradley & Willett, 2004; Vasilash, 1993; Melnyk et al., 1998).

In addition to noting the ways in which Kaizen events conform to the definition of an organizational project, it is almost important to identify key ways in which Kaizen events are likely to differ from "typical" organizational projects. These differences were kept in mind when reviewing the literature and developing the working theory of Kaizen event research – e.g., factors that relate to the longer-term coordination of project activities were omitted from the research. Primary differences include:

- *Project timeframe* – One week versus several months
- *Project scope* – As compared to many projects within organizations, Kaizen events have a relatively narrow scope. Scope is generally related to the length of the project timeframe. However, Kaizen events may have much shorter timeframes than, but similar scopes to, "traditional" CPI projects.

- *Need for monitoring and control during the project* -- While Kaizen events require planning up-front, due to their short timeframe and small scope, they do not require as much sophisticated tracking and control during the project; in fact, a lot of the traditional project management competences related to monitoring and control would not apply, or would apply only in a very simplified fashion, to the execution of Kaizen events.

Three models proposed in the project management literature are reviewed here and analyzed with respect to their apparent applicability to the study of Kaizen events. First, the early work of Pinto and Slevin (1987) is reviewed, since this work represents one of the first attempts to empirically identify the critical success factors for projects. Next, the more recent work of Belassi and Tukel (1996) and Nicolini (2002) are reviewed.

Building on prior, theoretical research, Pinto and Slevin (1987) sought to empirically identify the key input factors contributing to project success. Using a projective technique called "Project Echo," Pinto and Slevin asked a sample of 52 part-time MBA students at the University of Pittsburgh to imagine themselves in the role of project manager for a project in which they had recently been involved, and then to identify five actions they believed they could take to improve the likelihood of project success. Students were allowed to define "project success" however they chose; the only prompt provided was that a successful project was one that "resulted in organizational change" (Pinto & Slevin, 1987, p. 24). A total of 236 responses were collected. Two experts independently sorted the responses into categories – i.e., factors. Both experts identified 10 factors. However, the interrater agreement for individual responses was only 0.50 (Pinto & Slevin, 1987). These success factors, as described in a second article (Slevin & Pinto, 1986, p. 57-58) are: 1) *Project Mission* – i.e., goal clarity, 2) *Top Management Support*, 3) *Project Schedule/Plan* – i.e., the soundness and clarity of the tactical plan for achieving project objectives, 4) *Client Consultation*, 5) *Personnel*, 6) *Technical Tasks* – i.e., adequacy of technical knowledge and equipment for achieving the project mission, 7) *Client Acceptance*, 8) *Monitoring and Feedback*, 9) *Communication*, and 10) *Trouble-Shooting*.

Pinto and Slevin also developed an assessment tool, the Project Implementation Profile (PIP), a 100-item questionnaire intended to allow managers to rate the relative presence of the 10 critical success factors (Slevin & Pinto, 1986), and went on to conduct additional research on the relationship of the success factors to project outcomes for various types of projects (e.g., Pinto & Slevin, 1989; Pinto & Mantel, 1990). However, the PIP did not appear to be a reliable measurement tool for the research. Besides the apparent lack of rigorous testing – e.g., factor analysis – it is very long, with 10 items per scale, and several items have wording problems – e.g., double-barreled questions, etc. Pinto and Slevin's methods have many weaknesses – for instance, surveying students instead of actual project managers. The only requirement for participation was that students were employed full-time and had been part of at least one project team within their organizations within the two years prior to the study. Other weaknesses include low interrater agreement on classifications, and lack of rigorous testing for the PIP. However, their research is particularly interesting because it represents one of the first attempts to empirically identify project success factors. In addition, their critical success factor list is intuitively appealing, and many of these factors have been mentioned in subsequent studies (e.g., Belassi & Tukel, 1996; Nicolini, 2002), as well as descriptions from the practitioner literature – for instance, the Kaizen event literature mentions top management support (Vasilash, 1993; Rusiniak, 1996; Adams et al., 1997; "Keys to Success," 1997; Sheridan, 1997b; Vasilash, 1997; Cuscela, 1998; Hasek, 2000; Bicheno, 2001; Bane, 2002 Bradley & Willett, 2004; Martin, 2004) and clear project goals/objectives (Vasilash, 1993; Melnyk et al., 1998; Bradley & Willett, 2004; Martin, 2004) as potential success factors.

To develop their model of the relationship between critical success factors and project outcomes, Belassi and Tukel (1996) reviewed previous theoretical and empirical studies on project success factors, including the work of Pinto and Slevin (1987). Belassi and Tukel argue that many earlier studies focused on identifying lists of individual critical success factors, that were often either: 1) too general, or 2) too specific to particular types of projects. Instead of focusing on identifying individual critical success factors, Belassi and Tukel proposed a model focused on the relationships between four comprehensive types of critical success factors and project outcomes: 1) *factors related to the project* – e.g., size, uniqueness, urgency; 2) *factors related to the project team* – i.e., project manager and project team member characteristics; 3) *factors related to the organization* – e.g., top management support; and 4) *factors related to the external environment* – e.g., political, economic, social, technological, etc

In the Belassi and Tukel model, many variables traditionally thought of as critical success factors – such as the availability of resources, and the quality of project planning, scheduling and communication – are portrayed as intervening "system response" variables that are proposed to mediate the relationship between critical success factors – i.e., organization, environment, project team and/or project characteristics – and project outcomes. Belassi and Tukel tested the usefulness of their model by conducting a survey of project managers across a variety of industries – e.g., manufacturing, construction, defense, utilities. Managers were selected from the Project Management Institute directory. A total of 200 survey questionnaires were mailed, and 57 completed surveys were returned, for a response rate of 28%. The questionnaire consisted of two sections. The first section asked questions about the respondent's industry, typical project size, organizational structure, and the most important success criteria for the respondent's projects. The second section asked the respondents to identify all factors they believed were most important to successful project completion. Example factors were presented, grouped according to the four categories identified by Belassi and Tukel. In addition, each respondent was asked to specify, by category, additional factors he/she felt were most important to project success, which had not been included in the survey. Belassi and Tukel found that the results of this survey were useful for identifying differences in process success factors across industries, project size, organizational structures and success criteria. Belassi and Tukel found that factors related to project manager and team member characteristics, as well as the environment and the organization, had an impact on project success that differed by industry, project size, success criteria and organization type. In contrast, an earlier study by Tukel and another colleague (described in Tukel & Rom, 1998), which asked project managers from the same types of industries to identify the most important critical success factors for their projects, produced much less useful information. In the previous studies, project managers had been presented with a list of five critical success factors from the project management literature – i.e., top management support, client consultation, preliminary estimates, availability of resources, and project managers' performance – and asked to identify others that they felt were important, which were not part of these five. Study results indicated that organizational factors – e.g., top management support and the availability of resources – were the most highly rated factors across all types of industries, organizational structures, success criteria, and project sizes. Respondents rated project manager performance as much less important to project success than the project management literature suggested, and the free-specified responses failed to produce any useful categorizations. Belassi and Tukel suggest that their follow-up study – e.g., the one used to test the model – produced more useful information because their model was useful for helping project managers understand the interaction between categories of factors and project outcomes. Belassi and Tukel suggest that presenting the factors in a systematic way, by category rather than ungrouped list, helped the project managers provide more insightful and comprehensive responses about the relationships between critical success factors and project outcomes. The four categories identified by Belassi and Tukel, as well as their proposed interrelationships, were considered in the research model specification.

Nicolini (2002) proposed a provisional theoretical model describing the relationship between technical factors, social factors, and "project success." Nicolini's particular focus was on the social factors related to project success for construction industry projects, since the detailed study of social factors in projects had been neglected in the project management literature in general and the construction industry literature in particular. To develop the model, Nicolini paired a "top-down" theory-building approach – e.g., review of related literature streams – with a "bottom-up" approach – e.g. focus groups and interviews with construction industry practitioners – because there was no specific literature in the construction industry on the social factors related to project success. In the "top down" approach, Nicolini reviewed relevant work in organizational climate literature, cross-functional new product development team literature, and project management literature to identify technical and social factors that could contribute to successful team member interactions – Nicolini refers to this as good "project chemistry" – and to propose how these interactions could, in turn, influence project outcomes. For instance, the cross-functional team literature identifies five categories of factors related to the effectiveness of cross-functional teams, which Nicolini hypothesized may also be related to construction project success: 1) task design, 2) group composition, 3) organizational context, 4) internal processes and boundary management, and 5) group psychosocial traits.

In the "bottom-up" approach, Nicolini conducted two focus groups with a total of 17 participants and seven semi-structured interviews to elicit construction industry practitioners views. Three broad questions were addressed in both the focus groups and interviews: 1) *How do you describe project chemistry?*; 2) *What factors affect project chemistry?*; 3) *How does project chemistry affect project performance?* Interview and focus group data were transcribed verbatim and analyzed using qualitative content analysis procedures. Some construction industry practitioners defined "project chemistry" in terms of team member behaviors – e.g., open communication, collaboration, etc. – while others focused more on team member perceptions of the team environment – e.g., psychological safety; shared focus, etc. However, both clearly related to the outcomes of team member interactions. Similarly to the findings from his "top-down" literature review, Nicolini found that the construction industry practitioners believed that a variety of technical and social factors affected project chemistry, which in turn was believed to impact project performance, in terms of final product quality, project lead-time and project cost. Five categories of influential factors emerged from the focus group and interview data: 1) *commercial and business practice and task design* – this includes methods of establishing a relationship with the client and the quality of the product design process; 2) *team selection and composition* – this includes the team selection procedures – i.e., selecting the "right people" in terms of skill mix and personality – as well as the extent to which a stable team composition was maintained across one or several projects; 3) *quality of leadership* – this includes the effectiveness of the team leader; 4) *management of team development processes* – this includes the use of team development exercises, role negotiation, and the establishment of metrics related to both technical and social team outcomes; and 5) *initiatives to sustain involvement* – this includes communicating results to stakeholders throughout the larger organization and the client organization.

Nicolini used the results of both approaches to develop a model describing the proposed relationships between technical and social antecedent variables, team interactions – i.e., "project chemistry" – technical and social intermediate project outcomes, and, ultimately, "project success." Antecedent variables are divided into the variables that vary at the individual project level, called "Project Level Antecedents," and those that are likely to be "fixed" across all the projects in a given organization, called "Business Environment and Organizational Antecedents." "Project Level Antecedents" include team selection, team development processes, quality of leadership, etc. "Business Environment and Organizational Antecedents" include corporate Human Resource policies and the organization's product development process, etc. The outcomes of team member interactions – e.g. "project chemistry" – is measured by three dimensions from the team climate for innovation (TCI) model (Anderson & West, 1996, 1998): 1) clarity and level of agreement on objectives and vision; 2) quality and type of interaction; and 3) participative safety and mutual influence. The factor groups and interrelationship suggested by Nicolini were considered in the specification of the model for the research.

2.2.2 Team Effectiveness Theory

Cross-functional teams are the human resource structure of Kaizen events (Vasilash, 1993; Rusiniak, 1996; Adams et al., 1997; Sheridan, 1997b; Cuscela, 1998; Melnyk et al., 1998; LeBlanc, 1999; McNichols et al., 1999; Demers, 2002; Smith, 2003; Drickhamer, 2004b; Martin, 2004). Thus, team effectiveness theory can inform the study of Kaizen events, by identifying the types of input factors and process factors that may be relevant to the study of Kaizen events. The following section provides a brief review of the critical success factors identified in the team effectiveness literature, as identified in the comprehensive review provided in Cohen and Bailey (1997).

Cohen and Bailey reviewed research published between January 1990 and April 1996 on relationships between input and process factors and team effectiveness for work, parallel, project and management teams. The types of teams most relevant to the study of Kaizen events are parallel teams and project teams. Parallel teams are cross-functional teams focused on a specific task or problem that regular organizational structures are not well equipped to handle. As traditionally used in organizations, parallel teams often do not have authority to immediately implement changes but instead only have the power to recommend changes to senior management. "Traditional" CPI teams are often parallel teams. On the other hand, project teams are time-limited, cross-functional teams focused on a one-time product. The output of project teams may be an incremental improvement over an existing design or idea, or a radically new design or idea. In contrast to "traditional" CPI teams, Kaizen events more closely fit the definition of a project team, in that they are clearly time-limited, focused on a one-time product, and typically have considerable implementation authority.

Based on the trends in team research, Cohen and Bailey propose a heuristic model of team effectiveness, which describes interrelationships between key categories of "critical success factors" and team outcomes – e.g., effectiveness. *Environmental factors* represent broader industry, social, economic and political characteristics over which the organization has no direct control. Meanwhile, *design factors* are those characteristics of the task, group (e.g., team) or organization, which are ultimately under management influence. For instance, *task design* factors include the amount of autonomy given to the team for managing their work processes, as well as the interdependence of team member tasks. *Group design factors* include the size and demographic profile of the team. *Organizational design factors* include rewards, training and resources. *Group processes* describe the interactions between team members, as well as their interactions with external stakeholders. *Group psychosocial traits* describe group-level characteristics of the team, such as shared beliefs and emotions. Finally, *team effectiveness* can be measured through performance outcomes – such as the quality and quantity of the product/service provided, attitudinal outcomes – such as employee satisfaction and commitment, and behavioral outcomes – such as turnover and safety measures. The Cohen and Bailey model departs from a strict "input-process-output" approach to understanding team effectiveness (McGrath, 1984), by depicting a direct relationship between design factors i.e., inputs – and team effectiveness, as well as an indirect relationship via group process. Furthermore, group psychosocial traits, an intermediate outcome of design factors, is shown to be both directly related to team effectiveness and indirectly related to team effectiveness through group process.

2.2.3 Broader OSU – VT Research Initiative to Understand Kaizen Events

One additional, pre-existing framework was consulted in the specification of the research model: a preliminary, global research framework developed during the pilot phase of the broader OSU - VT research initiative to understand Kaizen events. This broader initiative, which received funding as a three year study from the National Science Foundation (NSF) in 2005, includes the current research and also a second research initiative focused on the study of the sustainability of Kaizen event results. As described in Farris, Van Aken, Doolen and Worley (2004), pilot efforts in this broader research initiative led to the specification of a preliminary global research framework, which describes the types of factors hypothesized to be related to both initial outcomes and sustainability of outcomes, as well as the types of outcomes to be studied. This framework was specified primarily using a "bottom-up" approach – i.e.., a preliminary investigation of the practitioner literature and direct observation of Kaizen events during the pilot effort – but did, less formally, incorporate some of the principal investigators' previous knowledge of team effectiveness and organizational change theory, as well as some preliminary review of the I/O psychology, organizational change and continuous improvement literature. This global research framework partitions relevant variables into: 1) input factors related to the target work area (*work area factors*); 2) input factors related to the Kaizen event design (*Kaizen event design factors*); 3) initial social system outcomes; 4) initial technical system outcomes; 5) longitudinal social system outcomes; and 6) longitudinal technical system outcomes. Longitudinal outcomes are intended to be measured at roughly six-month intervals over the period of 18 – 24 months. *Work area factors* and *Kaizen event design factors* are hypothesized to impact Kaizen event outcomes directly. There are no proposed, intervening process variables. In addition, there are proposed feedback loops between outcomes and work area factors and Kaizen event design factors, indicating that some Kaizen event design factors and work area factors may change over time as a result of event outcomes, thus potentially affecting sustainability of outcomes.

The model developed in the current research represents an extension of the preliminary global research framework, particularly in adding hypothesized intervening process variables proposed to mediate the relationship between event inputs and event outcomes, and in specifying the exact set of input and process variables expected to impact initial outcomes and the exact set of outcome variables to be studied. However, the preliminary global research framework provided an additional starting framework for specifying the model for the current research, by identifying two types of input factors – *Kaizen event design factors* and *work area factors* – which may be related to outcomes, as well as by stressing the need to measure both social and technical system outcomes.

2.2.4 Critical Success Factors from the Kaizen Literature

In addition to the "top-down" approach used to identify potential input and process factors from project management and team effectiveness theory, the researcher also used a "bottom-up" approach to analyze the Kaizen event literature to identify input and process factors potentially related to Kaizen event outcomes. The main purpose of this review was to summarize the event input and event process factors mentioned in the Kaizen event literature as describing Kaizen events and/or contributing to Kaizen event outcomes.

Methods used to identify references include: 1) ABI/Inform database searches for terms such as "Kaizen event" and "kaizen blitz;" 2) searches for the same terms using Google, the Virginia Tech library catalog and amazon.com; 3) dialogue with committee members and other researchers; and 4) review of the literature review section, if it existed, for all references identified. However, most references did not include a literature review on Kaizen events.

While reviewing each Kaizen event reference, the researcher took careful note of: 1) any input or process factor mentioned; and 2) example benefits – i.e., technical and social system outcomes – mentioned. As the researcher took notes, a list of potential event input and process factors was iteratively compiled and organized. To facilitate understanding and summarization of the factors, a factor categorization scheme was developed from the first 33 sources reviewed. First, the input and process factors were initially compiled into an uncategorized list (Appendix A), while noting which factors were mentioned by multiples sources. The factors were then inductively grouped into categories the researcher developed (see Appendix B). This was an iterative process involving an initial grouping and a refinement that combined some initial groups. Finally, the Pinto and Slevin model, Belassi and Tukel model, Nicolini model and Cohen and Bailey model were analyzed to develop categories of input and process factors. Based on this analysis, the researcher developed five categories of factors related to: 1) task design, 2) team design, 3) organizational support, 4) event process, and 5) the broader context of the events – i.e., characteristics of the Kaizen event program, such as the strategic coordination of multiple Kaizen events. These final categories of factors identified are presented in Appendix C. It is important to note that the factors in the fifth category – i.e., those related to the broader context of the events -- are outside the scope of the research, although they would be relevant for a follow-up study with a higher-level focus – i.e., a study of the design and aggregate effects of Kaizen event programs within organizations. Some of these factors were measured to describe the context of the current research, but were not studied as independent variables in this research.

As more references beyond the initial 33 were identified and reviewed, factors mentioned in these sources have been mapped to the initial categories. Most recently, in May 2005, a comprehensive ABI/Inform search for the more general term "kaizen" was performed. This search identified 360 potential sources. The abstracts for all sources were reviewed to determine whether the source appeared to be focused on short-term Kaizen events – e.g., "kaizen blitz," "kaizen projects," etc. – or longer-term kaizen activities – e.g., CPI. Sources which focused on short-term Kaizen events and articles for which the focus was not clear were fully reviewed. A total of 55 references, including the initial 33, have been reviewed and incorporated to date. Table 1 shows the current list of factors identified from the Kaizen event literature, by category.

The existing Kaizen event literature is largely practitioner-focused and anecdotal. In particular, the existing literature lacks systematic, explanatory studies of Kaizen events across multiple organizations. Most of the references are articles that reflect companies' first-hand experience with Kaizen events. These articles generally focus on describing Kaizen events within a single organization, often discussing the Kaizen event program in general, rather than describing individual Kaizen events in detail, although brief examples of the process and/or results of individual events are often included. The articles usually include anecdotal accounts of the types of technical and social system benefits achieved – although social system benefits are rarely measured, as reported in Kosandal & Farris, 2004 – as well as factors that article authors felt were relevant for describing the design and operation of the company's Kaizen events. Sometimes, Kaizen events are the main focus of the articles. Other times, Kaizen events are mentioned in a broader article on successful or innovative practices used by a particular company. Many of the articles occur in non peer-reviewed trade publications (e.g., *Industry Week*, *Works Management*, etc.). However, some occur in peer-reviewed publications (e.g., *The TQM Magazine,* etc.) but have a similar focus to articles appearing in non peer-reviewed publications -- describing Kaizen events, presenting example results from Kaizen events in one or a few companies, and providing suggestions for designing Kaizen events based on the author's observations or suggestions from corporate representatives, rather than on systematic research. Other references – in particular, three guidebooks (Laraia et al., 1999; Mika, 2002; Vitalo et al., 2003) – prescribe guidelines for organizations interested in conducting Kaizen events, but lack research support for these guidelines.

Only a minority of the references found clearly addressed a research audience (Adams et al., 1997; Melnyk et al., 1998; Patel et al., 2001; Patil 2003; Bradley & Willett, 2004; Kumar & Harms, 2004; Bateman, 2005). However, these references still primarily concentrated on describing Kaizen events as a change mechanism and the apparent results from Kaizen events within one or a few companies. None of the references found contained systematic, explanatory studies of multiple Kaizen events across multiple organizations.

For instance, Adams, Schroer and Stewart (1997) present a case study of the use of an integrated improvement methodology ("Quick-Step™ process improvement"), which includes the use of kaizen events, within a single organization. The case study includes a description of the Quick-Step™ methodology, as well as aggregate results achieved between 1993 and 1995. Although Quick-Step™ includes other elements, one key element is 28- or 48-hour team-based "kaizen sessions" – i.e., Kaizen events – where teams implement "as many improvements as possible immediately" (Adams et al., 1997, p. 26). These kaizen sessions are not the main focus on the article, which concentrates on the overall Quick-Step™ methodology and aggregate results achieved. Thus, the article does not presented a detailed description of the process and/or results of individual kaizen sessions. However, Adams, Schroer and Stewart (1997) indicate that kaizen sessions have been used to achieve both technical system and social system results – e.g., increased action orientation and waste awareness. Furthermore, the article does not represent a designed, explanatory study of process improvement methodology effectiveness – i.e., testing of success factors – but rather the description of a specific improvement methodology.

Melnyk, Calantone, Montabon and Smith (1998) describe key features of Kaizen events, such as Kaizen event objectives, process and tools – e.g., takt time analysis and 5S are briefly described. In addition, they provide some example results from one firm using Kaizen events, including aggregate improvements from the Kaizen event program as a whole -- e.g.., in cycle time, WIP, floor space, etc. – and sample results from two individual Kaizen events. All the results presented relate to technical system outcomes, although Melnyk, Calantone, Montabon and Smith (1998) state that the case example firm was using Kaizen events to help create a learning organization – i.e., a social system outcome. Melnyk, Calantone, Montabon and Smith (1998) also briefly relate the development of Kaizen events to organizations' past experience with other change mechanisms, including traditional continuous improvement and business process reengineering (BPR). The Melnyk, Calantone, Montabon and Smith (1998) article is not a designed study of event effectiveness – i.e., it is not an explanatory study that empirically tests the impact of pre-selected event input and process factors. In addition, the Melnyk, Calantone, Montabon and Smith (1998) work only presents example results from a single organization. However, this article is clearly intended for a research audience. Melnyk, Calantone, Montabon and Smith (1998) state that their article was intended to: 1) create an awareness of Kaizen events as a change mechanism, and 2) create a research interest in Kaizen events, because they were previously un-researched as a change mechanism.

In a case study of four manufacturing companies, Patel, Dale and Shaw (2001) mention that "kaizen teams" were being used by at least one of the four organizations to implement set-up time reduction. However, Patel et al. do not provide any description of the Kaizen event process or results. Therefore, it is not completely clear that these “kaizen teams” represent Kaizen events, although pilot work for the Oregon State University – Virginia Tech initiative to understand Kaizen events, as well as the literature review for the research, suggest that many organizations use Kaizen events for set-up time reduction (e.g., Minton, 1998; Bradley & Willett, 2004)

In an unpublished master’s thesis, Patil (2003) conducted a field study of a single Kaizen event within the Boeing 757 manufacturing facility. The Kaizen event chosen was the first event conducted in a particular work area – section 44 of the 757 program, which manufactures part of the fuselage. The focus of this study was to determine whether the results from the event were sustained and, if not, to suggest potential reasons for the lack of sustainability. Patil used the Kaizen event action plan to develop a sustainability checklist. The target work area was audited in August 2003, approximately eight months after the initial Kaizen event (December 2002). The audit consisted of direct observation by Patil and interviews with target work area employees. It was concluded that the event results were not sustained. During the interviews, Patil also used a “5 Why” probe to identify potential reasons for the lack of sustainability. Finally, fishbone diagramming using the results of the interviews and Patil’s observations of the target area were used to suggest reasons for the lack of sustainability, as well as to suggest guidelines for sustaining Kaizen event results. Patil calls these guidelines a “standard framework” for sustaining Kaizen events. The guidelines developed through this process are: 1) involving employees in Kaizen event planning and decision making; 2) promising job security to employees; 3) training employees in the benefits of kaizen; 4) training employees in Kaizen event process management; 5) establishing a clear link between organizational strategy and Kaizen events; 6) establishing standard operating procedures (SOPs), based on the changes made during events; 7) creating mechanisms to review the implementation progress of Kaizen event teams – i.e., progress on completing action items post-event; 8) establishing an internal communication network for sharing training and performance information; 9) developing a cross-functional team for Kaizen events; 10) proving support for employees – e.g., time to complete action items after the event; 11) clearly identifying boundaries for jobs in SOPs — i.e., “to do” versus “not to do” instructions; and 12) routinely identifying employee needs using an employee survey. In addition, Patil conducted a survey of employees in the target area to test for differences in perceptions between two job categories (mechanics versus managers) on content areas such as: the relationship between company strategy and Kaizen events, Kaizen event process management, team working and learning, training and education, management support, and communication. All survey questions were developed by Patil, and it is not clear what, if any, previous work or theory they are based on. The Patil work suffers from several methodological weaknesses including: an overall lack of detail with respect to the methods and measures used; lack of justification for the case selection -- e.g., why did Patil select the first Kaizen event in the target work area?; lack of justification for the measures used – e.g., how were the survey items developed? what previous research are they based on?; weaknesses of the measurement instruments – e.g., many of the survey questions were double-barreled – "My team leader's commitment and involvement is excellent;" and, finally, lack of justification for the conclusions drawn. The conclusions drawn were based solely on Patil's subjective interpretation of the largely qualitative data gathered. While the conclusions drawn appear to be logical, they lack the support that would be provided by statistical analysis of a larger sample, with more refined measures. However, the Patil work can inform future research as an exploratory work in the area of Kaizen event sustainability.

Bradley and Willett (2004) is another work targeted at a research audience. This article describes an educational partnership between Cornell University and a local company, Lord Corporation, that allowed business and industrial engineering students to participate in Kaizen events first-hand. The article provides both a description of the educational partnership and a description of Lord Corp.'s lean manufacturing program. In addition, the article describes the goals and objectives, and timeframes of the 12 Kaizen events – Bradley and Willett refer to them as "kaizen projects" – in which the Cornell students participated. Finally, Bradley and Willett describe "success factors" which Cornell students identified in "debriefing sessions" as contributing to the success or failure of the events in which they participated. In addition, Bradley and Willett identify the factors that Lord Corp.'s managers reported as key to Kaizen event success, although Bradley and Willett do not indicate what methods were used to gather this information. While the main purpose of the article is to show how the innovative educational partnership between university and industry can be beneficial to both sides, the description of Lord Corp's lean manufacturing program is useful from a research perspective, and the "success factors" identified are intuitively appealing and may be useful for both researchers and practitioners. However, the article is entirely qualitative, while lacking the depth and systematic purpose that would be expected from a designed case study focused entirely on Kaizen events, rather than on describing the outcomes of the educational partnership. The "success factors" have not been tested in any fashion, although they do provide a useful starting point for future research.

Kumar and Harms (2004) present a case study of process improvement within one manufacturing firm, Viratec Thin Films, Inc. Kaizen events are not the focus of the article, which focuses on more general process improvement strategies. However, Kaizen events are presented as one of two process improvement vehicles used by the organization. The other vehicle is process mapping workshops. To describe the role of Kaizen events in process improvement, Kumar and Harms describe the organization's first Kaizen event, which occurred in a manufacturing process, including a description of the event process and technical systems outcomes. The description also includes comments from participating employees representing social system outcomes – e.g., enjoyment of Kaizen events, learning about how to design manufacturing processes, etc. The article suggests that Kaizen events can be a useful vehicle for identifying and implementing process improvements, but this rests primarily on the description of one Kaizen event. Kumar and Harms state that the organization continued to use Kaizen events to achieve further improvement. Since Kaizen events were not the focus of the article, the event described does not represent a detailed, designed study – i.e., a formal case study or field study – but rather a post-hoc description used for illustrative purposes. However, this work is clearly focused at a research audience interested in strategies for process improvement.

Finally, Bateman (2005) focuses on identifying factors related to the sustainability of process improvements (PI). While not specifically described as such, the "PI workshops" studied in Bateman (2005) appear to be Kaizen events. They are described as "hands-on" improvement activities that incorporate lean concepts, and they last five days. Bateman (2005) does not attempt to identify any factors related to initial event outcomes. However, Bateman makes several contributions to the study of the sustainability of event outcomes. First, Bateman describes a framework – originally proposed in Bateman and David (2002) – for classifying overall levels of sustainability. This framework was developed based on the observation that companies often do not collect accurate longitudinal data on Kaizen event outcomes, so it is not always possible to calculate the exact percentage of improvements sustained. In the Bateman and David (2002) framework, an "A" level of sustainability occurs when all PI workshop improvements are sustained, both those made initially and those made by completing follow-up action items) and further improvement is made. A "B" level of sustainability occurs when all PI workshop improvements are sustained but no further improvement is made. A "C" level occurs when all improvements made initially are sustained but action-items are not closed out, and a "D" level occurs in the reverse situation. An "E" level describes a workshop that fails to substantially sustain any improvements. Using this framework, 40 PI workshops in 21 organizations were analyzed to determine which organization-level and work area-level factors were related to the two highest categories of sustainability – i.e., "A" and "B" levels. The organization-level and work area-level factors were measured through structured interviews, with close-ended questions using a binary response scale – i.e., "Yes" if the factor was present and "No" if the factor was absent. Bateman (2005) found that PI workshops that fully sustained all outcomes – i.e., "A" and "B" levels of sustainability – occurred in organizations where time was allocated to 5S each day and in work areas where performance was regularly measured and reviewed, there was a formal process for shop floor employees to communicate problems to management, work area managers supported PI and shop floor employees had a high degree of decision-making authority in their daily work. In addition, "A" levels of sustainability occurred in organizations with a full-time PI coordinator, senior management support for PI, formal processes for training work area employees who were not on the PI team in the new work methods, and a clear link between organization-level strategy and work area strategy. Furthermore, "A" levels of sustainability occurred in organizations where work area managers became directly involved in PI activities.

Table 1. Factor Groups for Kaizen Event Factors from the Kaizen Literature

1. **Task** -- Task Design Factors (Cohen & Bailey, 1997); Project Level Antecedents (Nicolini, 2002); Factors Related to Project (Belassi & Tukel, 1996)

   a) Event Duration
      - One week or shorter (*LeBlanc, 1999; Oakeson, 1997; Vasilash, 1997; Drickhamer, 2004b; Watson, 2002; Smith, 2003; Cuscela, 1998; McNichols et al., 1999; Martin, 2004; Sheridan, 1997b; Patton, 1997; Bradley & Willett, 2004; Vasilash, 1993; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998; Sheridan, 2000a; Heard, 1997; Pritchard, 2002; Klaus, 1998; Clark, 2004; David, 2000; Wittenberg, 1994; "Get Smart, Get Lean," 2003; Sabatini, 2000; Tanner & Roncarti, 1994; Larson, 1998a; "Waste Reduction Program Slims Fleetwood Down," 2000; Kumar & Harms, 2004; Gregory, 2003; Taylor & Ramsey, 1993*)
      - Two weeks or shorter (*Minton, 1998; Demers, 2002; Harvey, 2004; Foreman & Vargas, 1999*)

   b) Team Authority
      - Teams have implementation authority (*LeBlanc, 1999; Oakeson, 1997; Minton, 1998; Martin, 2004; Sheridan, 1997b; Bradley & Willett, 2004; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998; Pritchard, 2002; Klaus, 1998; Wheatley, 1998; Laraia, 1998; "Get Smart, Get Lean," 2003; Sabatini, 2000; Tanner & Roncarti, 1994; Larson, 1998a; Treece, 1993; "Waste Reduction Program Slims Fleetwood Down," 2000; Kumar & Harms, 2004; Taylor & Ramsey, 1993; Foreman & Vargas, 1999*)
      - Team controls starting and stopping times of Kaizen event activities – often, long days 12-14 hrs (*Sheridan, 1997b; Vasilash, 1993; Larson, 1998b; Tanner & Roncarti, 1994; Kumar & Harms, 2004*)
      - Team members participate in setting improvement goals and assigning team roles (*Heard, 1997*)
      - Team has considerable control over the activities they adopt in meeting event goals (*Wheatley, 1998; Larson, 1998a; Tanner & Roncarti, 1994*)
      - Team identifies own improvement opportunities and targets (*Wittenberg, 1994*)
      - Team appoints own leader (*Wittenberg, 1994*)
      - Team leader participates in setting goals (*Tanner & Roncarti, 1994*)
      - Problem scope can be shrunk or expanded during the Kaizen event (*Tanner & Roncarti, 1994*)
      - Team selects target area (*Kumar & Harms, 2004*)

   c) Problem Scope
      - Require a standard, reliable target process/work area as input (*LeBlanc, 1999; Bradley & Willett, 2004*)
      - Requires a well-defined problem statement as input (*Rusiniak, 1996; Adams et al., 1997*)
      - Avoid problems that are too big and/or emotionally involved (*Rusiniak, 1996; Sheridan, 1997b; "Get Smart, Get Lean," 2003; Gregory, 2003*)
      - Preference given to Kaizen events that require simple, well-known tools versus more complex tools (*Bradley & Willett, 2004*)
      - Avoid problems that require advanced statistical analysis (*Harvey, 2004*)

- Can be used for process design (*Harvey, 2004*)

d) Event Goals
- Linked to organizational strategy (*LeBlanc, 1999; "Keys to Success," 1997; Melnyk et al., 1998; "Get Smart, Get Lean," 2003*)
- Challenging "stretch" goals (*LeBlanc, 1999; Minton, 1998; Rusiniak, 1996; Cuscela, 1998; Bradley & Willett, 2004; Bicheno, 2001; Tanner & Roncarti, 1994; Treece, 1993; Kumar & Harms, 2004; Gregory, 2003*)
- Focused – on a specific process, product, or problem (*Minton, 1998; Drickhamer, 2004b; Martin, 2004; Sheridan, 1997b; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998; Heard, 1997; David, 2000; Laraia, 1998; Tanner & Roncarti, 1994*)
- Concrete, measurable goals (*Martin, 2004; Bradley & Willett, 2004; Vasilash, 1993; Melnyk et al., 1998; Heard, 1997; Treece, 1993; "Waste Reduction Program Slims Fleetwood Down," 2000; Foreman & Vargas, 1999*)
- Used to implement lean manufacturing (*Vasilash, 1997*)
- Kaizen events are focused on the needs of the external customer – e.g. improving value – versus internal efficiency (*Melnyk et al., 1998; Laraia, 1998*)
- Kaizen events are focused on waste elimination (*Watson, 2002; Cuscela, 1998; Martin, 2004; Patton, 1997; Adams et al., 1997; Heard, 1997; Klaus, 1998; David, 2000; Wittenberg, 1994; "Get Smart, Get Lean," 2003; Tanner & Roncarti, 1994; "Waste Reduction Program Slims Fleetwood Down," 2000; Kumar & Harms, 2004; Taylor & Ramsey, 1993; Foreman & Vargas, 1999*)
- Kaizen events are focused on improving flow (*Tanner & Roncarti, 1994*)

2. **Team** -- Group Composition Factors (Cohen & Bailey, 1997); Project Level Antecedents (Nicolini, 2002); Factors Related to the Project Team (Belassi & Tukel, 1996)

a) Team size
- 3 – 5 people (*Rusiniak, 1996*)
- 6 – 10 people (*McNichols et al., 1999; Martin, 2004; Vasilash, 1993; Laraia, 1998; Sabatini, 2000; Tanner & Roncarti, 1994*)
- 10 – 12 people (*LeBlanc, 1999; Demers, 2002; Watson, 2002; Pritchard, 2002; Laraia, 1998; Sabatini, 2000; Larson, 1998b; Treece, 1993*)
- 12 – 14 people (*Cuscela, 1998; Laraia, 1998; Larson, 1998b; Treece, 1993*)
- 14 - 15 people (*Laraia, 1998; Larson, 1998b*)

b) Use of Cross-Functional Teams (*LeBlanc, 1999; Drickhamer, 2004b; Rusiniak, 1996; Demers, 2002; Smith, 2003; Cuscela, 1998; McNichols et al., 1999; Martin, 2004; Sheridan, 1997b; Vasilash, 1993; Vasilash, 1993; Adams et al., 1997; Melnyk et al., 1998; Sheridan, 2000a; Pritchard, 2002; Laraia, 1998; Harvey, 2004; Foreman & Vargas, 1999*)
- Team Structure
  - Informal "floating" team structure (*Adams et al., 1997*)
  - Team members volunteer to participate (*Watson, 2002; Adams et al., 1997*)
  - Team leader and sub-team leader are selected by the business unit manager (*Tanner & Roncarti, 1994*)
- Functional Heterogeneity
  - Including "fresh eyes" – people with no prior knowledge of the target area – on the team (*LeBlanc, 1999; Vasilash, 1997; Kleinsasser, 2003; Minton, 1998; Cuscela, 1998; McNichols et al., 1999; Martin, 2004; Bradley & Willett, 2004; Melnyk et al., 1998; David, 2000; Foreman & Vargas, 1999*)
  - Including people from the work area on the Kaizen event team (*Redding, 1996; Minton, 1998; Womack & Jones, 1996a; Martin, 2004; Sheridan, 1997b; Bradley & Willett, 2004; Vasilash, 1993; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998; Heard, 1997; David, 2000; Wheatley, 1998; Tanner & Roncarti, 1994; Treece, 1993; Taylor & Ramsey, 1993*)
  - Most team members are from work area (*Tanner & Roncarti, 1994*)
  - Including people from all production shifts in Kaizen event team (*Vasilash, 1993*)
  - Each team member has specific knowledge of the process (*Watson, 2002*)
  - Each team member is either directly or indirectly involved in the target process (*Kumar & Harms,*

2004)

- Including people from all functions required to implement/sustain results on the Kaizen event team (*Bradley & Willett, 2004; Vasilash, 1993; Adams et al., 1997*)
- Including subject matter experts (SMEs) – e.g., quality engineers. Maintenance – on the team (*David, 2000; Treece, 1993; Taylor & Ramsey, 1993*)
- Including only one employee per department on the Kaizen event team, except for the department being blitzed, to avoid over-burdening any department (*Minton, 1998*)
- Including managers and supervisors on the Kaizen event team (*Oakeson, 1997; "Keys to Success," 1997; Vasilash, 1993; Bicheno, 2001; Heard, 1997; Clark, 2004; David, 2000; "Get Smart, Get Lean," 2003; Sabatini, 2000; Tanner & Roncarti, 1994; Treece, 1993; Taylor & Ramsey, 1993)*
- Including target area supervisor on Kaizen event team *(Patton, 1997*)
- Including customers on the Kaizen event team (*Hasek, 2000; Vasilash, 1997; McNichols et al., 1999; Vasilash, 1993; Adams et al., 1997; Melnyk et al., 1998; Heard, 1997; Larson, 1998b; Treece, 1993*)
- Including suppliers on the Kaizen event team (*Vasilash, 1997; McNichols et al., 1999; Vasilash, 1993; Adams et al., 1997; Melnyk et al., 1998; Heard, 1997; "Get Smart, Get Lean," 2003; Tanner & Roncarti, 1994; Larson, 1998b*)
- Including benchmarking partners or other external non-supply chain parties on the Kaizen event team (*McNichols et al., 1999; Sheridan, 1997b; Vasilash, 1993; "Get Smart, Get Lean," 2003*)
- Including people from other sister plants or corporate headquarters on the team (*Sabatini, 2000; Tanner & Roncarti, 1994*)
- Avoid including people from competing plants or functions on the Kaizen event team (*Bradley & Willett, 2004*)

- Team Member Problem-Solving Abilities
  - Black Belts assigned to Kaizen event teams for Lean-Six Sigma programs (*Sheridan, 2000b*)
  - At least one member of Kaizen event team experienced enough in tool(s) to teach others (*Bradley & Willett, 2004*)
  - Including outside consultants on the Kaizen event team, particularly for the first few Kaizen events (*Oakeson, 1997; Bicheno, 2001; Sabatini, 2000; "Waste Reduction Program Slims Fleetwood Down," 2000; Kumar & Harms, 2004*)
- Team Member Attitudes
  - Team members are positive thinkers (*"Get Smart, Get Lean," 2003*)

3. **Organization** -- Organizational Context Factors (Cohen & Bailey, 1997); Project Level Antecedents (Nicolini, 2002); Factors Related to the Organization (Belassi & Tukel, 1996)

a) Management Support/Buy-In (*Bane, 2002; Hasek, 2000; Vasilash, 1997; Rusiniak, 1996; Cuscela, 1998; Martin, 2004; Sheridan, 1997b; "Keys to Success," 1997; Bradley & Willett, 2004; Vasilash, 1993; Bicheno, 2001; Adams et al., 1997; Heard, 1997; Laraia, 1998; Tanner & Roncarti, 1994; Treece, 1993; "Waste Reduction Program Slims Fleetwood Down," 2000; Kumar & Harms, 2004; Taylor & Ramsey, 1993*)

- Plant manufacturing director temporarily moves his/her office to Kaizen event room during event (*Tanner & Roncarti, 1994*)
- Business unit managers divide their time between the shop floor and the Kaizen event room during the event (*Tanner & Roncarti, 1994*)

b) Resource Support

- Team members dedicated only to Kaizen event during its duration (*Minton, 1998; McNichols et al., 1999; Martin, 2004; Bradley & Willett, 2004; Bicheno, 2001; Melnyk et al., 1998; Heard, 1997; Harvey, 2004; Kumar & Harms, 2004; Gregory, 2003; Foreman & Vargas, 1999*)
- Having support personnel – e.g., maintenance, engineering, etc. – "on call" during the event, to provide support as needed – e.g., moving equipment overnight (*McNichols et al., 1999; Martin, 2004; Sheridan, 1997b; Bradley & Willett, 2004; Bicheno, 2001; Adams et al., 1997; Wittenberg, 1994; Tanner & Roncarti, 1994; Gregory, 2003; Taylor & Ramsey, 1993*)
- Low cost solutions (*Purdum, 2004; Cuscela, 1998; Martin, 2004; Sheridan, 1997b; Vasilash, 1993;*

Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998; Klaus, 1998; Tanner & Roncarti, 1994; Larson, 1998a; Treece, 1993; Taylor & Ramsey, 1993)

- Cost is <u>not</u> a factor (*Minton, 1998*)
- Dedicated room for Kaizen event team meetings (*Creswell, 2001; Tanner & Roncarti, 1994*)
- Snacks provided to team during Kaizen event (*Creswell, 2001; Adams et al., 1997*)
- Stopping production in target area during the Kaizen event (*Bradley & Willett, 2004*)
- Priority given to Kaizen team requests (*Kumar & Harms, 2004*)
- Use of a Kaizen cart containing tools and supplies that serves as a mobile office for the Kaizen event team during the event (*Taylor & Ramsey, 1993*)

c) Rewards/Recognition

- Rewards and recognition for team after the event – e.g., celebrations (*Adams et al., 1997; Melnyk et al., 1998; Martin, 2004; Tanner & Roncarti, 1994; Larson, 1998b; Taylor & Ramsey, 1993; Foreman & Vargas, 1999*)

d) Communication

- Importance of buy-in from employees in work area (*Sheridan, 1997b*)
- Kaizen event team members from work area encouraged to discuss event activities and changes with others in the work area during the event to create buy-in (*Bicheno, 2001*)
- Discussion of changes with employees in the work area during the Kaizen event (*Wittenberg, 1994; Sabatini, 2000; Gregory, 2003*)

e) Event Planning Process

- Well-defined and thorough event planning activities – i.e., adequate preparation (*Sheridan, 1997b; Bradley & Willett, 2004; Heard, 1997; Gregory, 2003; Foreman & Vargas, 1999*)
- Including process documentation – e.g., VSM, process flowcharts, videotapes of the process, current state data, etc. – as input to Kaizen event (*Minton, 1998; McNichols et al., 1999; Martin, 2004; Bradley & Willett, 2004; Bicheno, 2001; David, 2000; Kumar & Harms, 2004; Gregory, 2003*)
- Notifying employees in adjoining work areas before the start of the Kaizen event – e.g., publicizing the event (*McNichols et al., 1999; Gregory, 2003*)
- Use of a Kaizen mandate – e.g., Kaizen event announcement – to clearly define and communicate event goals (*Heard, 1997; Foreman & Vargas, 1999*)
- Tools/problem solving method to be used are identified by the facilitator (*Heard, 1997*)
- Team leader prepares a briefing package with historical performance data, layout drawings, staffing data and customer requirements data before the event, which is given to the rest of the team on the first day of the event (*Tanner & Roncarti, 1994*)
- Development of an "event schedule" – i.e., a high-level road map of activities – before the event (*Foreman & Vargas, 1999*)

f) Training

- Less than two hours of formal training provided to team (*Minton, 1998; McNichols et al., 1999)*
- Including ½ day of training at the start of the event – i.e., training in tools, kaizen philosophy, etc. (*Vasilash, 1993; Melnyk et al., 1998; Heard, 1997; Klaus, 1998; David, 2000; Tanner & Roncarti, 1994; Treece, 1993; "Waste Reduction Program Slims Fleetwood Down," 2000; Foreman & Vargas, 1999*)
- Including 1 day of training at the start of the event – i.e., training in tools, kaizen philosophy, etc. (*Wittenberg, 1994; "Get Smart, Get Lean," 2003; Larson, 1998b; "Waste Reduction Program Slims Fleetwood Down," 2000; Taylor & Ramsey, 1993; Foreman & Vargas, 1999*)
- Facilitators provide "short courses" on topics "on the spot" if a team gets stuck (*Minton, 1998*)
- Team members who aren't from the process get training in the process and may even work in the production line for a few days before the Kaizen event (*Minton, 1998*)
- Including ergonomics training as part of Kaizen event training (*Wilson, 2005*)
- Including "team-building" exercises as part of Kaizen event training (*Bicheno, 2001; Foreman & Vargas, 1999*)
- Making sure that each participant has thorough knowledge of the "seven wastes" prior to team activities (*Bicheno, 2001*)
- Training can be provided before the formal start of the event – i.e., offline (*McNichols et al., 1999; Bicheno, 2001; Gregory, 2003*)

4. **Event Process** -- Internal Process Factors (Cohen & Bailey, 1997); Processes (Nicolini, 2002); Project Manager's Performance on the Job (Belassi & Tukel, 1996)

   a) Action Orientation (*LeBlanc, 1999; Redding, 1996; Smith, 2003; Martin, 2004; Sheridan, 1997b; Patton, 1997; Vasilash, 1993; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998*; *Sabatini, 2000; Tanner & Roncarti, 1994; Larson, 1998a; Treece, 1993; Taylor & Ramsey, 1993; Foreman & Vargas, 1999*)
      - Involve first-hand observation of target area – e.g., data collection, etc. (*Smith, 2003; Vasilash, 1993; Clark, 2004; David, 2000; Wittenberg, 1994; Tanner & Roncarti, 1994; Larson, 1998b; Treece, 1993; "Waste Reduction Program Slims Fleetwood Down," 2000; Kumar & Harms, 2004; Taylor & Ramsey, 1993; Foreman & Vargas, 1999*)
      - Keep line running during Kaizen event, because it is important for the team to observe a running line (*Sheridan, 1997b; Sabatini, 2000; Larson, 1998a; Tanner & Roncarti, 1994; Kumar & Harms, 2004*)
      - Cycles of solution refinement during Kaizen event (*Bradley & Willett, 2004; Bicheno, 2001; Melnyk et al., 1998; Clark, 2004; "Waste Reduction Program Slims Fleetwood Down," 2000; Taylor & Ramsey, 1993*)
      - Training work area in employees in the new process is part of the Kaizen event (*Martin, 2004; Heard, 1997*)
   b) Problem Solving Tools/Techniques
      - Videotapes of setups (*Minton, 1998; Bradley & Willett, 2004*)
      - Brainstorming (*Minton, 1998; Watson, 2002; Martin, 2004; Bradley & Willett, 2004; Vasilash, 1993; Pritchard, 2002; Laraia, 1998; Kumar & Harms, 2004; Taylor & Ramsey, 1993)*
      - Avoid preconceived solutions (*Rusiniak, 1996; Bradley & Willett, 2004*)
      - Seek improvement, not optimization (*Rusiniak, 1996; Vasilash, 1993*)
      - Question the current process – ask why things are done the way they are (*Watson, 2002; Minton, 1998; Taylor & Ramsey, 1993*)
      - Team should not be too rigid about sticking to formal methodology (*Bradley & Willett, 2004*)
      - Creating a video report-out (*Sabatini, 2000*)
      - Decisions are driven by hard/quantitative data (*Tanner & Roncarti, 1994; Gregory, 2003*)
      - Tools used depend on event goals – e.g., SMED, 5S, etc. (*Tanner & Roncarti, 1994*)
   c) Team Coordination
      - At least one member of Kaizen event team keeps the team "on track" – i.e., focused (*Bradley & Willett, 2004; Vasilash, 1993; Wheatley, 1998; Foreman & Vargas, 1999*)
      - Use of subteams (*Minton, 1998; McNichols et al., 1999; Sheridan, 1997b; Bicheno, 2001; Sabatini, 2000; Treece, 1993; Foreman & Vargas, 1999*)
      - Use of a Kaizen newspaper/30-day action item list to capture needed actions that cannot be implemented during the Kaizen event (*"Winning with Kaizen," 2002; McNichols et al., 1999; Martin, 2004; Bradley & Willett, 2004; Melnyk et al., 1998; Heard, 1997; Larson, 1998a; Treece, 1993; Tanner & Roncarti, 1994; Gregory, 2003*)
      - Team reviews current progress to plan next day's activities (*Wheatley, 1998; Sabatini, 2000*)
      - Every 2 – 3 hours, team reassembles in Kaizen event room to review progress and then returns to the target work area (*Tanner & Roncarti, 1994*)
      - Posting team actions, metrics, concepts and data around the team meeting room during the event (*Foreman & Vargas, 1999*)
      - Kaizen event team gives daily updates to management, where managers hear the team's plans and give input (*Foreman & Vargas, 1999*)
   d) Participation
      - Involving everyone on the Kaizen event team in the solution process (*Vasilash, 1993*)
      - Making each team member responsible for implementing at least one improvement idea (*Bicheno, 2001*)
      - Each team member participates in report-out to management (*Adams et al., 1997; Larson, 1998b*)

5. **Broader Context** (Kaizen Event Program Characteristics)

a) Kaizen Event Deployment
   - Spacing out events – e.g., only one event per quarter (*Taninecz, 1997)*
   - Concurrent Kaizen events (*Vasilash, 1997; Watson, 2002; Cuscela, 1998; Bradley & Willett, 2004; Adams et al., 1997; Wittenberg, 1994; Tanner & Roncarti, 1994; Gregory, 2003*)
   - Targeted at areas that can provide a "big win" – i.e., provide a big impact on the organization (*Minton, 1998; Cuscela, 1998; Martin, 2004; Sheridan, 1997b; "Keys to Success," 1997; Bradley & Willett, 2004; Melnyk et al., 1998; Tanner & Roncarti, 1994; Gregory, 2003*)
   - Repeat Kaizen events in a given work area (*"Winning with Kaizen," 2002; Purdum, 2004; Womack & Jones, 1996a; McNichols et al., 1999; Sheridan, 1997b; Bradley & Willett, 2004; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998*)
   - Can be held based on employee suggestions for improvement (*Jusko, 2004; Watson, 2002; Heard, 1997*)
   - Kaizen events used in non-manufacturing areas – e.g., office Kaizen events (*Womack & Jones, 1996a; Sheridan, 1997b; Bradley & Willett, 2004; Melnyk et al., 1998; Klaus, 1998; Baker, 2005; Clark, 2004; Foreman & Vargas, 1999*)
   - Combining Kaizen events with other improvement approaches (*Bicheno, 2001*)
   - Using a sequence of related Kaizen events – e.g., 5S, SMED, Standard Work – to progressively improvement a given work area (*Bicheno, 2001; Melnyk et al., 1998; Laraia, 1998; Treece, 1993*)
   - Attack "low hanging fruit" (*Smith, 2003; Bicheno, 2001; Heard, 1997; Clark, 2004*)
   - Output of given Kaizen event is used to determine the next Kaizen event (*Adams et al., 1997*)
   - Using Kaizen events sparingly, as method of achieving breakthrough change and overturning current paradigms (*Sheridan, 2000a*)
   - Holding Kaizen events across different areas of the organization or value stream (*Heard, 1997*)
   - Daily team leader meetings for concurrent Kaizen events (*Wittenberg, 1994; Sabatini, 2000; Tanner & Roncarti, 1994*)
   - First Kaizen event targeted at highest volume, most important product (*Larson, 1998a*)
   - Use of shorter, informal or "mini" Kaizen events (*Tanner & Roncarti, 1994; "Waste Reduction Program Slims Fleetwood Down," 2000*)
   - Concurrent Kaizen event teams are co-located – i.e., share the same meeting room (*Tanner & Roncarti, 1994*)
   - Using Kaizen events to address areas of concern in value stream maps (VSM) (*Gregory, 2003*)
   - Concurrent Kaizen event teams brief each other two times each event day (*Gregory, 2003*)

b) Organizational Policies/Procedures
   - "No layoffs" policy (*Redding, 1996; Vasilash, 1997; Creswell, 2001; "Winning with Kaizen," 2002; Womack & Jones, 1996a; "Keys to Success," 1997; Bradley & Willett, 2004; Melnyk et al., 1998; Larson, 1998a; Treece, 1993; Tanner & Roncarti, 1994; "Waste Reduction Program Slims Fleetwood Down," 2000*)
   - Organization-wide commitment to change (*Redding, 1996*)
   - Total alignment of organizational procedures and policies with Kaizen event program (*"Keys to Success," 1997; Tanner & Roncarti, 1994*)
   - Organization-wide communication of the philosophies behind and importance of Kaizen events (*Kumar & Harms, 2004*)

c) Kaizen Program Support
   - Use of a "Kaizen office," including full-time coordinators/facilitators *(Heard, 1997; "Keys to Success," 1997; Bicheno, 2001; Foreman & Vargas, 1999*)
   - Keeping a central database of employee Kaizen event participation, past Kaizen event results, ideas for future Kaizen events, training materials, etc. (*Heard, 1997*)
   - 30 days sustainability reviews for Kaizen events (*Heard, 1997*)
   - Offline training in new processes for employees not trained during the event – i.e., second shift, etc. (*Heard, 1997*)
   - Use of a consultant to get the Kaizen event program started – i.e., to help set up the Kaizen event promotion office, etc. (*Heard, 1997; Martin, 2004*)

- Emphasis on follow-up – e.g., consultants stayed on for 5 –10 days after the event to help standardize achievements (*Kumar & Harms, 2004; Gregory, 2003*)
- Follow-up Kaizen events with "traditional" kaizen (CPI) activities (*Gregory, 2003*)
- Kaizen event team continues to meet regularly after the event to track open action items (*Foreman & Vargas, 1999*)

2.3 Research Model Specification

Justification for measuring both technical system and social system outcomes is provided in the Kaizen event literature, STS theory and project management and team effectiveness theory. As will be described in more detail in the next chapter, the technical system measures chosen for study represent both objective performance measures and perceptual measures. The social system measures chosen for study include a range of employee KSAs that are aligned with continuous improvement. These variables were chosen to reflect some of the major human resource benefits cited in the Kaizen event practitioner literature. The KSA framework is an established framework from the I/O psychology literature (Muchinsky, 2000).

Overall, the specification of the event input and event process factors to be studied in the research was more difficult than the specification of outcome measures. Due to the large number of potential factors in the Kaizen event literature, as well as project management and team effectiveness theory, it was necessary to identify a smaller set of key variables for study. A major goal of this refining process was to identify at least one factor related to each of the four relevant factor groups identified in the review of the Kaizen event literature: 1) task design; 2) team design; 3) organizational support; and 4) event process. Definitions of the specified variables were provided in Chapter 1. However the following paragraphs provide detail on the specification of event input and event process factor.

In the specification of initial research model, task and team design factors (see Table 1) are grouped together as *Kaizen Event Design Antecedents*, since both sets of factors describe aspects of the design of a given Kaizen event – e.g., its goals, team composition, etc. The Kaizen event design antecedent factors chosen for study in the research are:

- *Goal Clarity* – A task design factor that reflects the clarity of event goal characteristics (see Table 1).
- *Goal Difficulty* – A task design factor that reflects team perceptions of goal difficulty (see Table 1).
- *Team Kaizen Experience* – A team design factor that reflects team compositional characteristics (see Table 1) – specifically, the experience of Kaizen team members with Kaizen events.

- *Team Functional Heterogeneity* – A team design factor that reflects team compositional characteristics (see Table 1) – specifically, the diversity of functional expertise for Kaizen team members.
- *Team Autonomy* – A task design factor that reflects Kaizen event team authority (see Table 1).
- *Team Leader Experience* – While not generally emphasized in the Kaizen event literature, this team design factor is emphasized in the Nicolini model (2002), Belassi and Tukel model (1996), and practitioner-oriented texts on designing Kaizen events (Mika, 2002).

Organizational and Work Area Antecedents include organizational factors identified in the Kaizen event practitioner literature (see Table 1), as well as characteristics of the target work area. Including characteristics of the work area, as well as the organization, is important since both appear to be drivers of Kaizen event design antecedents, as well as the Kaizen event process and outcomes. For instance, the complexity of the target system could directly influence team composition, as well as team activities. The organizational and work area antecedents selected for study in the research are:

- *Management Support* – An organizational design factor (see Table 1). Management support may also contain certain aspects of rewards/recognition – another organizational design factor (see Table 1). For instance, the event budget may contain funds for a team celebratory lunch immediately following the event.
- *Event Planning Process* – An organizational design factor (see Table 1). Event planning process may also contain aspects of communication — another organizational design factor (see Table 1). For instance, planning may include a meeting with work area employees to announce the event and gain buy-in.
- *Work Area Routineness* – A work area factor (see Figure 5) that also relates to event scope (see Table 1).

Kaizen Event Process Factors include process factors from Table 1. The factors in this category are:

- *Action Orientation* (see Table 1) – This particular variable was also chosen for study since it is one of the distinguishing factors of Kaizen events – e.g., one of the most frequently cited differences between Kaizen events and "traditional" CPI activities. In addition, team effectiveness theory also suggests that this variable could have an important impact on Kaizen event outcomes. *Action Orientation* also reflects aspects of team coordination, since it describes how the team managed its time.
- *Affective Commitment to Change* – While this appeared to be perceived more as an input factor than as a process factor in the Kaizen event literature, the Nicolini (2002) model and the Cohen and Bailey (1997) model

suggest that this factor is more properly categorized as a process factor – e.g., a factor arising from the Kaizen event design, organizational and work area antecedents, and/or team activities.

- *Tool Appropriateness* (see Table 1) – Based on the Kaizen event literature and pilot research, it appears that most Kaizen event teams use some form of structured problem solving tools and techniques – e.g., brainstorming, spaghetti diagramming, SMED, etc. However, the tools/techniques used appear to vary by the type event goals – e.g., setup reduction events use SMED, while standard work events use spaghetti diagramming, etc. Thus the selection and use of appropriate problem solving tools is expected to be an important factor in Kaizen event team effectiveness. In addition to collecting ratings on tool appropriateness, the current research also collected a list of problem-solving tools used. Thus, differences based on tools used could be investigated in future post-hoc analysis.
- *Tool Quality* (see Table 1) – Quality of tool use must be measured separately from tool appropriateness, since it would be possible for a team to make a poor selection of tools, but do a good job actually applying the tools, or to select the right tools but do a poor job of applying them. Both scenarios – and any in between – could be expected to have different effects on event outcomes.
- *Internal Processes* – this construct relates both to Team Coordination and Participation (see Table 1). It describes the extent to which team interactions were harmonious, including open communication and respect for each individual's contribution.

One variable that was not ultimately included in the current research, but may be of interest in future research is *Training* (see Table 1). The current research classified *Training* as an organizational and work area antecedent, since it is a precursor to event problem-solving activities and may not even be conducted for the given event, if all team members have participated in similar events. However, *Training* could also be considered part of the event process. Based on the practitioner literature and the pilot research, it seems likely that all organizations provide training for new members of the Kaizen event team. However, this training may be provided "offline" if other members of the Kaizen event team have previously participated in similar events (McNichols et al., 1999; Bicheno, 2001). Thus, the binary variable of whether or not team members ultimately receive some form training is not likely to vary across events or organizations. However, whether the training occurs as part of the event, the length of training, topics covered and perceived effectiveness of training may vary across events and/or organizations. In the current research, contextual information on training length, training topics and whether training was conducted as a formal part of the event was collected through the Team Activities Log and the Event Information Sheet. Some of these factors could therefore be investigated in future post-hoc analysis. In addition, each organization's general approach to Kaizen event training was conducted through the interview data describing the organizations overall approach to conducting Kaizen events. However, collection of additional data related to training – such as the perceived adequacy of training by Kaizen event team members, could be of interest in future research. As will be discussed more in Chapter 6, studying the perceived adequacy of training may be difficult, since it could easily be confounded with employee perceptions of the overall outcomes of the event on their KSAs.

CHAPTER 3: RESEARCH METHODS

The following sections describe the research design in terms of the ways in which the factors of interest were measured, the way the data were collected, and the way the data were prepared for hypothesis testing – i.e., data screening and initial data analyses to describe the factor structure, confirm scale reliability and support aggregation. The general research design is a multi-site field study using a cross-sectional design. Kaizen events within study organizations were sampled and measures were taken on the factors of interest, allowing the statistical analysis of the relationships between event input factors, event process factors, technical system outcomes and social system outcomes – i.e., the empirical testing of the working theory of Kaizen event effectiveness (see Chapter 1 and Chapter 2).

3.1 Operationalized Measures for Study Factors

The following section summarizes the variables to be studied in the research. For each event input factor, event process factor, and outcome, an operationalized measure was developed. Definitions of these measures were provided in Chapter 1. The following sections describe the input data collected to calculate each measure, the instrument used to collect the input data, the measurement timing and the data source – i.e., Kaizen event team members or the event facilitator. The operationalized measures represent a mixture of objective and perceptual measures. Operationalized measures were developed using factor descriptions from the literature review. Wherever possible, survey questionnaire measures were based on existing survey scales. However, actual item wording was modified to reflect the specific context of the current research. In addition, the first two events studied in this research were considered a pilot phase and study instruments and methods were analyzed and refined based on this additional pilot testing.

Two data sources were used to collect the data – the Kaizen event team members and the Kaizen event facilitator. The Kaizen event facilitator is the individual who coordinates event planning and provides guidance to the team during the event. Often, the Kaizen event facilitator is a member of management, a technical expert or a person who facilitates events as their full-time job. The facilitator is not considered a part of the event team, but rather acts as a support resource or coach. The facilitator typically delivers training to the team on the first day of the event, and may provide guidance on how to use tools during the event and/or help keep team discussions "on track." The facilitator also helps the team procure needed resources, including meeting space, equipment, approval for changes, etc. However, event decisions are made by team members and typically one of the team members acts as the team leader.

3.1.1 Operationalized Measures for Technical System Outcomes

The operationalized measures for the technical system outcomes are: *% of Goals Met*, *Impact on Area* and *Overall Perceived Success.* While *% of Goals Met* captures the objective impact of Kaizen events – e.g., event success relative to goals, *Impact on Area* and *Overall Perceived Success* capture stakeholder perceptions of event success. Stakeholder perceptions of *Impact on Area*, while likely primarily focused on the extent of work area, process or product improvements, may also capture more subjective improvement dimensions – e.g., perceptions of an improved work area climate, etc. Similarly, stakeholder perceptions of *Overall Perceived Success* may be related to performance versus goals, impact on the target system, and perceptions of the amount of buy-in from management and other personnel. One goal of this research was to investigate the degree of correlation between the objective measure *% of Goals Met* and the perceptual measures. The operationalized measures are shown in Table 2.

Table 2. Operationalized Measures for Technical System Outcomes

Variable	Input Data	Measurement Instrument	Measurement Timing	Data Source
1. % of Goals Met	• *team improvement goals* • *post-event performance on goals* • *relative importance of goals* – main goal versus secondary goal *% of Goals Met* is computed as the average % of main goals met	Event Information Sheet	Following the report-out meeting	Facilitator, Report Out File
2. Impact on Area (IMA)	Four-item scale based on the *Impact on Area* scale developed and assessed for reliability in pilot research (Doolen et al., 2003b); measured using a 6-point Likert response scale: • IMA1: "This Kaizen event had a positive effect on this work area." • IMA2: "This work area improved measurably as a result of this Kaizen event." • IMA3: "This Kaizen event has improved the performance of this work area." • IMA4: "Overall, this Kaizen event helped people in this area work together to improve performance."	Report Out Survey	Immediately following the report-out meeting	Team
3. Overall Perceived Success (OVER)	Single-item based on a measure developed in pilot research (Farris et al., 2004); measured using a 6-point Likert response scale: • "Overall, this Kaizen event was a success"	Report Out Survey, Event Information Sheet	Following the report-out meeting	Team, Facilitator

3.1.2 Operationalized Measures for Social System Outcomes

The operationalized measures for the social system outcomes are shown in Table 3. The social system outcome scales were based on scales developed by Doolen et al. (2003b), using the KSA framework from the I/O psychology literature (Muchinsky, 2000). All three dimensions – i.e., knowledge, skills, and attitudes – describe employee characteristics that are required to adequately perform desired tasks – in this case, continuous improvement activities. "Knowledge" describes the body of information necessary, "skills" refer to psychomotor capabilities, and "attitudes" refer to cognitive capabilities – e.g., desire to perform the given activity. In all survey scale measures used in this research, a group-level referent – i.e., "our team" – rather than an individual-level referent – e.g., "I," was used. The research assumes a referent-shift model of team composition (Chan, 1998). This model assumes the variables of interest – e.g., management support, team autonomy and KSAs – occur at the group-level rather than the individual-level. Although data are still collected at the individual-level, because it is impossible to directly collect perceptual data at the group-level, the unit of interest is the group, thus the referent is the group. Provided the actual data statistically support aggregation, team member averages are then used as the variable measure for each team. It is important to note here that organizational learning theory suggests that KSAs related to specific Kaizen events occur at the group level, rather than the individual level. Kaizen events include the collective interpretation and collective action on problems, which are characteristics of the group, rather than individual learning process (Crossan et al., 1999). In addition, Kaizen events contain the three dimensions of group learning identified by Groesbeck (2001) in field research on group learning – experimenting, collaborating and integrating the group's work with the larger organization. Thus, it is theoretically sound to posit that increases in KSAs – i.e., learning – if they occur, occur at the group, rather than individual, level in Kaizen event teams.

Table 3. Operationalized Measures for Social System Outcomes

Variable	**Input Data**	**Measurement Instrument**	**Measurement Timing**	**Data Source**
1. Understanding of CI (UCI)	Four-item survey scale based on the *Understanding of Need for Change* and *Understanding of Need for Kaizen* scales developed for pilot research by Doolen et al. (2003b); measured using a 6-point Likert response scale: • UCI1: "Overall, this Kaizen event increased our team members' knowledge of what continuous improvement is." • UC2: "In general, this Kaizen event increased our team members' knowledge of how continuous improvement can be	Report Out Survey	Immediately following the report-out meeting	Team

	applied." • UCI3: "Overall, this Kaizen event increased our team members' knowledge of the need for continuous improvement." • UCI4: "In general, this Kaizen event increased our team members' knowledge of our role in continuous improvement."			
2. Attitude (AT)	Four-item scale based on the *Attitude* scale developed and assessed for reliability in pilot research (Doolen et al., 2003b); measured using a 6-point Likert response scale: • AT1: "In general, this Kaizen event motivated the members of our team to perform better." • AT2: "Most of our team members liked being part of this Kaizen event." • AT3: "Overall, this Kaizen event increased our team members' interest in our work." • AT4: "Most members of our team would like to be part of Kaizen events in the future."	Report Out Survey	Immediately following the report-out meeting	Team
3. Skills (SK)	Four-item scale based on the *Skills* scale developed and assessed for reliability in pilot research (Doolen et al., 2003b); measured using a 6-point Likert response scale. • SK1: "Most of our team members can communicate new ideas about improvements as a result of participation in this Kaizen event." • SK2: "Most of our Kaizen event team members are able to measure the impact of changes made to this work area." • SK3: "Most of our team members gained new skills as a result of participation in this Kaizen event." • SK4: "In general, our Kaizen event team members are comfortable working with others to identify improvements in this work area."	Report Out Survey	Immediately following the report-out meeting	Team

3.1.3 Operationalized Measures for Event Process Factors

The operationalized measures for the event process factors are shown in Table 4.

Table 4. Operationalized Measures for Event Process Factors

Variable	**Input Data**	**Measurement Instrument**	**Measurement Timing**	**Data Source**
1. Action Orientation (AO)	Four-item scale not based on a preexisting scale; measured using a 6-point Likert response scale: • AO1: "Our team spent as much time as possible in the work area." • AO2: "Our team spent very little time in our meeting room."	Report Out Survey	Immediately following the report-out meeting	Team

	• AO3: "Our team tried out changes to the work area right after we thought of them." • AO4: "Our team spent a lot of time discussing ideas before trying them out in the work area." (REVERSE CODED)			
2. Affective Commitment to Change (ACC)	Six-item scale based on the validated scale developed by Herscovitch and Meyer (2002); measured using a 6-point Likert response scale: • ACC1: "In general, members of our team believe in the value of this Kaizen event." • ACC2: "Most of our team members think that this Kaizen event is a good strategy for this work area." • ACC3: "In general, members of our team think that it is a mistake to hold this Kaizen event." (REVERSE CODED) • ACC4: "Most of our team members that this Kaizen event will serve an important purpose." • ACC5: "Most of our team members think that things will be better with this Kaizen event." • ACC6: "In general, members of our team believe that this Kaizen event is needed."	Kickoff Survey	Immediately following the kickoff meeting	Team
3. Tool Appropriateness	For each problem-solving tool used by the team, the facilitator was asked to rate the *appropriateness* of using the tool to address the team's goals using a 6-point Likert response scale. *Tool appropriateness* is calculated as the average *appropriateness* rating across all tools.	Event Information Sheet	Following the report-out meeting	Facilitator
4. Tool Quality	For each problem-solving tool used by the team, the facilitator was asked to rate the *quality* of the team's use of the tool using a 6-point Likert response scale. *Tool quality* is calculated as the average *quality* rating across all tools.	Event Information Sheet	Following the report-out meeting	Facilitator
5. Internal Processes (IP)	Five-item scale based on *Internal Processes* dimensions identified by Hyatt and Ruddy (1997); measured using a 6-point Likert response scale: • IP1: "Our team communicated openly." • IP2: "Our team valued each member's unique contributions." • IP3: "Our team respected each others' opinions." • IP4: "Our team respected each others' feelings." • IP5: "Our team valued the diversity in our team members."	Report Out Survey	Immediately following the report-out meeting	Team

3.1.4 Operationalized Measures for Kaizen Event Design Antecedents

The operationalized measures for the kaizen event design antecedents are shown in Table 5.

Table 5. Operationalized Measures for Kaizen Event Design Antecedents

Variable	Input Data	Measurement Instrument	Measurement Timing	Data Source
1. Goal Clarity (GC)	Four-item scale based on the scale developed by Wilson, Van Aken and Frazier (1998); measured using a 6-point Likert response scale. "Our team spent as much time as possible in the work area." • GC1: "Our team has clearly defined goals." • GC2: "The performance targets our team must achieve to fulfill our goals are clear." • GC3: "Our goals clearly define what is expected of our team." • GC4: "Our entire team understands our goals."	Kickoff Survey	Immediately following the kickoff meeting	Team
2. Goal Difficulty (GDF)	Four-item scale. The first and third items are based on the *Goal Difficulty* scale developed by Ivancevich and McMahon (1977) and adapted by Hart, Moncrief and Parasuraman (1989). The other two items are not based on a preexisting scale; measured using a 6-point Likert response scale: • GDF1: "Our team's improvement goals are difficult." • GDF2: "Meeting our team's improvement goals will be tough." • GDF3: "It will take a lot of skill to achieve our team's improvement goals." • GDF4: "It will be hard to improve this work area enough to achieve team's goals."	Kickoff Survey	Immediately following the kickoff meeting	Team
3. Team Kaizen Experience	*Previous Kaizen Event Experience* -- the number of previous Kaizen events in which each team member has participated. *Team Kaizen Experience* is computed as the average number of previous Kaizen events per team member.	Kickoff Survey and Report Out Survey	Immediately following the kickoff meeting and the report out meeting, respectively	Team
4. Team Functional Heterogeneity	*Functional Area* – the job function of each team member – e.g., "operator," "technician," "engineer," "supervisor," "manager," "other" – as reported by the facilitator. *Team Functional Heterogeneity* is measured by an index of variation for categorical data, *H* (Shannon, 1948), as reported in Teachman (1980). This index has also been used in research on group diversity (e.g., Jehn et al., 1999; Pelled et al., 1999; Jehn & Bezrukova, 2004).	Event Information Sheet	Following the report-out meeting	Facilitator

	$H = \sum_i p_i(\log(1/p_i))$			
5. Team Autonomy (TA)	Four-item survey scale. The first two items were based on the *Group Autonomy* scale developed and validated by Kirkman and Rosen (1999) and further revised by Groesbeck (2001). The last two items were based on the *Employee Empowerment* scale created by Hayes (1994); measured using a 6-point Likert response scale: • TA1: "Our team had a lot of freedom in determining what changes to make to this work area." • TA2: "Our team had a lot of freedom in determining how to improve this work area." • TA3: "Our team was free to make changes to the work area as soon as we thought of them." • TA4: "Our team had a lot of freedom in determining how we spent our time during the event."	Report Out Survey	Immediately following the report-out meeting	Team
6. Team Leader Experience	The number of previous Kaizen events that the team leader has led or co-led	Event Information Sheet	Following the report-out meeting	Facilitator

3.1.5 Operationalized Measures for Organizational and Work Area Antecedents

The operationalized measures for the organizational and work area antecedents are shown in Table 6

Table 6. Operationalized Measures for Organizational and Work Area Antecedents

Variable	Input Data	Measurement Instrument	Measurement Timing	Data Source
1. Management Support (MS)	Five-item scale. The first two items were based on the *Resource Allocation* scale developed by Doolen et al. (2003a). The last three items are not based on a preexisting scale; measured using a 6-point Likert response scale. • MS1: "Our team had enough contact with management to get our work done." • MS2: "Our team had enough materials and supplies to get our work done." • MS3: "Our team had enough equipment to get our work done." • MS4: "Our team had enough help from our facilitator to get our work done." • MS5: "Our team had enough help from others in our organization to get our work done."	Report Out Survey	Immediately following the report-out meeting	Team
2. Event Planning Process	Total person-hours invested in preparing for the improvement event.	Event Information Sheet	Following the report-out meeting	Facilitator
3. Work Area Routineness	Composite measure of *product mix complexity* (i.e., stability of product mix) and *routinization* (i.e., degree to which the production flow is similar for a given product produced at different time intervals and for different products produced by the work area). These dimensions were identified based on: the technology classification framework by Perrow (1967), the environmental uncertainty framework by Duncan (1972), the product/volume – layout/flow production system classification matrix by Miltenberg (1995), and the task routinization scale developed by Withey, Daft and Cooper (1983) and used by Gibson and Vermeulen (2003). In particular, the first item is based on Withey, Daft and Cooper (1983). Measured using a 6-point Likert response scale. • WAC1: "The work the target work area does is routine." • WAC2: "The target work area produces the same product (SKU) most of the time." • WAC3: "A given product (SKU) requires the same processing steps each time it is produced." • WAC4: "Most of the products (SKUs) produced in the work area follow a very similar production process."	Event Information Sheet	Following the report-out meeting	Facilitator

3.2 Overview of Data Collection Instruments

Table 7 describes the instruments to be used to collect data on the operational variables, as well as the context of the event.

Table 7. Data Collection Activities for Each Event Studied

Instrument	**Variables Measured**	**Timing**	**Description**	**Data Source**
Kickoff Survey	• Goal Clarity • Goal Difficulty • Affective Commitment to Change • Team Kaizen Experience	Immediately following the kickoff meeting at the beginning of the Kaizen event	19 item survey questionnaire with cover page and instructions (see Appendix E)	Team
Team Activities Log	None directly – provides an understanding of event context and can be compared to *Action Orientation* scale results	One member of the Kaizen team completes the Team Activities Log during the event activities	Blank document with spaces for the team member to record the activities of the team as they occur (see Appendix F). Broken down by day with half hour intervals	Team
Report Out Survey	• Attitude • Skill • Understanding of CI • Impact on Area • Overall Perceived Success • Team Autonomy • Management Support • Action Orientation • Internal Processes • Team Kaizen Experience	Immediately following the report-out of team results at the end of the Kaizen event.	39 item survey questionnaire with cover page and instructions (see Appendix G)	Team
Event Information Sheet	• Team Leader Experience • Work Area Routineness • Event Planning Process • Overall Perceived Success • Team Functional Heterogeneity • Team Size • % of Goals Met • Tool Appropriateness • Tool Quality	Following the report-out meeting – target was one to two weeks after the event	15 item questionnaire with cover page and instructions (see Appendix H)	Facilitator

In addition to the data collection activities described in Table 7, the Kaizen event report out file created by the team was colleted for each event. This was used to provide back-up and contextual data on the event – e.g., team size/composition, team goals, *% of Goals Met*.

In addition to the collection of the data for each Kaizen event studied, the researcher collected the following data on basic organizational characteristics that may impact the generalizability of study results:

- *Basic organizational demographic data* – location, industry sector, major products, number of employees, number of local facilities.
- *Information on the history of the Kaizen event program* – e.g., how long the organization has been conducting Kaizen events, types of results experienced, difficulties experienced, basic methodology for conducting Kaizen events – selecting, planning, implementing and sustaining events, etc.

This second type of organizational information was collected using a semi-structured interview guide – the Kaizen Event Program Interview Guide (see Appendix I) – which was developed and pilot tested as part of the broader effort to understand Kaizen events.

3.3 Data Collection Procedures

3.3.1 Sample Selection

The final sample consisted of six organizations who were also participating in the broader OSU – VT initiative to understand Kaizen events. To provide some basis for comparison across organizations and to reduce unwanted – "nuisance" – variability across organizations, the following boundary conditions were used to select organizations to participate in this research:

- The organizations manufacture products of some type – e.g., no purely service or knowledge-work organizations were included. This ensured that organizations have some baselines similarities in focus, fundamental processes and metrics used to measure performance. However, organizations across different industries were recruited to increase generalizability of results.
- The organizations must have been conducting Kaizen events for at least one year prior to the start of the study. This criterion was intended to eliminate the organizational learning start-up curve from organizations that are just starting to implement Kaizen events. This allows the study of Kaizen events in organizations that are fairly "mature" in the use of events, providing a "best practice" sample for studying Kaizen event effectiveness. Additional research on Kaizen events within organizations just beginning to use this mechanism may be of interest in future research.

- The organizations must use Kaizen events systematically, as part of a formal organizational improvement strategy, rather than as "single use" change mechanisms. Again, this is intended to provide a "mature," "best practice" sample by including companies that embrace the philosophical and strategic roots of Kaizen events, rather than companies that view Kaizen events primarily as an ad-hoc problem-solving tool and only use them sporadically.
- The organizations must conduct Kaizen events relatively frequently – i.e., at least one event per month on average. This was intended to allow an adequate sample size of Kaizen events within each organization.

Through information gained from colleagues, industrial partners, conference presentations and trade/scholarly publications, the research team sought out organizations known to fit these boundary conditions, across a fairly broad spectrum of industries. These organizations were provided with a short description of the research and the expected benefits from participation, and were invited to participate. Thus, the sample at the organizational level was not randomly selected; however, the necessity of utilizing the boundary conditions made the selection of a random sample infeasible, if not impossible. However, participation in the overall OSU – VT study of Kaizen events is open to all organizations that fit the boundary conditions. While not part of the sample used in this research, additional calls for participation in the broader study were included in conference presentations and meetings, as well as in trade journals and in manufacturing and industrial engineering society newsletters. Two additional organizations have signed on to participate in future research that is part of the broader study.

Seven organizations originally agreed to participate. However, one organization withdrew after only providing data for one Kaizen event. Therefore, the final sample size was six at the organizational level. Table 8 summarizes the characteristics of the participating organizations.

Within each organization, Kaizen events were randomly selected for study. The original objective was to sample all the Kaizen events conducted within the sample timeframe – i.e., approximately January 2006 – August 2006. Three organizations – Company A, Company B and Company C – agreed to this sampling frequency. However, certain organizations requested a lower sampling frequency – Company D, Company E and Company F. In these organizations, a systematic sampling procedure was adopted (Scheaffer, et al., 1996). Where the average number of events in the company per month was some number n, a k was selected between one and n, such that every kth event was targeted for study. Overall, the actual frequency of events studied was generally lower than the target frequency of events studied, due to some cases of non-response from the organizations regarding their current event schedules and some cases of facilitators failing to administer surveys. Table 9 lists the estimated achieved versus targeted response rate for each company in terms of percentage of events studied.

Table 8. Characteristics of Study Organizations

	Company A	Company B	Company C	Company D	Company E	Company F
Description	Secondary wood product manufacturer	Electronic motor manufacturer	Secondary wood product manufacturer	Manufacturer of large transportation equipment	Specialty equipment manufacturer	Steel component manufacturer
First Kaizen Event	1998	2000	1992	1998	2000	1995
Kaizen Event Rate (end of 2005)	2 – 3 per month	About 1 every other month in steady state; however, every 6-12 months they also hold "umbrella events," where 5-7 events are held concurrently	2 per month	3- 4 per facilitator per month	2 per week	About 1 per month
% of Organization that has Experienced Events	100%	90%	Not sure	85%	100%	20%
Major Processes Targeted	Operations	Operations, sales/marketing, customer service/technical support, product design/redesign, production planning/inventory control, process design/redesign	Operations	Engineering and related activities	All areas of organization	Manufacturing, order entry, accounts receivable, distribution, vendors, engineering product development
% Manufacturing Events	Almost 100% manufacturing	75% manufacturing	Almost 100% manufacturing	70% non-manufacturing	Not sure	80-85% manufacturing

The main sampling criterion for including events was that the Kaizen event was considered a "formal event" by the organization and included all the basic support processes associated with Kaizen events – e.g., planning, formal announcement, report out, etc. This criterion was added since many organizations run shorter versions of Kaizen events -- often one or two days long – which are not considered full Kaizen events. These events are often spur of the moment and often do not contain all the elements associated with Kaizen events. For instance, the work in the event may be interrupted – i.e., two days off, two days on – or part-time versus full-time, the report out may be to a work area supervisor versus management, and training is often omitted. In addition, different tools may be used and different problems/goals addressed. Although these shorter informal events are interesting – and may be of particular interest in future research (i.e., comparing formal versus informal events) – they appeared to be different enough from formal events to avoid including them in this initial study.

Table 9. Estimated Response Rates from Study Organizations

Organization	# Events	% Studied	Target %	Study Window
Company A	15	100%	100%	October 2005 – May 2006
Company B	8	56%	100%	March 2006 – May 2006
Company C	11	100%	100%	January 2006 – June 2006
Company D	4	13%	25% - 33%	January 2006 – June 2006
Company E	12	33%	50%	January 2006 – June 2006
Company F	6	24%	60%	January 2006 – May 2006
Total Events	56			

Appendix R summaries the characteristics of the events studied in the participating organizations. In some cases, events could not be included due to missing data – i.e., a response rate lower than 50% on either the Kickoff Survey and/or Report Survey. In addition, following the factor analysis of the survey scales (described later in this chapter), a sixth team was excluded due to insufficient sample size at a variable. Table 10 summarizes the final number of events per company that were included in the study. A total of 51 events out of the original 56 sampled were included in the final analysis.

Table 10. Final Count of Events Included in the Study

Organization	# Events
Company A	15
Company B	8
Company C	7
Company D	4
Company E	11
Company F	6
Total Events	51

3.3.2 Mechanics of the Data Collection Procedures and Data Management

The following section provides an overview of the data management procedures. The author developed an Excel spreadsheet ("Data Collection Checksheet") to track the data collection activities for each event. Raw data were inputted into a second Excel spreadsheet. Completed surveys and other hard copies of data collection instruments were stored within a secure location at Virginia Tech – i.e. a file cabinet in the researcher's office. Electronic data were stored by the author on secure computers – i.e. the author's office computers provided and monitored by the ISE department, and a personally owned laptop computer.

The data collection procedures were designed to be as stand-alone as possible, both to avoid instrumentation error, and to facilitate the collection of data at remote sites. Because this research studied Kaizen events occurring concurrently at multiple organizations, many of which occurred on the west coast, it would not have been possible for the researcher to personally administer the survey questionnaires during team report-out meetings – because most organizations schedule report-out meetings for Fridays, this would have literally required the researcher to be in two places at once. In addition, building the data collection procedures into the normal Kaizen event process in the study organizations reduced the potential for bias from an external person attending the meetings – i.e., allowed for more naturalistic data. All data collection tools were designed to be self-administered The Kaizen event facilitators served as the data collection coordinators for their organizations, and were trained using standard instructions for administering and collecting the questionnaires (see Appendix Q for some of the administration and training tools used to train facilitators in the data collection process). The facilitators mailed completed Kickoff Surveys, Report Out Surveys and Team Activities Logs back to the OSU-VT research team at the end of the event. The Event Information Sheet was completed electronically. In addition, the OSU-VT research team created a secure website to allow organizational contacts – e.g., Kaizen event facilitators – to upload supporting files – e.g., Kaizen event announcements, report-outs – electronically.

One final note is that, because the research involved human subjects, IRB approval was needed. This approval was sought, and was received, as part of the broader OSU – VT research initiative. Participation was voluntary for all parties involved in the study. No compensation of any form was provided. However, participating organizations did receive research summaries describing results within their organization, as well as across organizations. It is not anticipated that individuals face any risk from participating in the research, and confidentiality of individual results has been carefully protected.

3.4 Data Screening

Kickoff Survey and Report Out Survey data were manually entered into Excel spreadsheets by the researcher and by a PhD candidate/graduate research assistant at OSU. The researcher entered all survey data for Companies A, B, C and D – i.e., the east coast companies – and the OSU student entered all survey data for Companies E and F – i.e., the west coast companies. The two researchers maintained separate master Excel spreadsheets and used the same basic data verification method. Data spreadsheets were regularly posted to the secure research website to be shared with other members of the research team.

Immediately after entry, each survey was checked against the spreadsheet at least once to verify the accuracy of the data entry. The researcher checked each survey two times, while the OSU student checked each survey once. Both researchers used the "find/replace" function in Excel to locate missing values and to replace them with a symbol to indicate that the cell value was missing versus accidentally deleted, etc. The researcher used a "-" to indicate a missing value, while the OSU student used a "N/A." Reverse coded variables were flagged and recoded using Excel formulas (i.e., 7-x). Construct averages for the initial scales were also calculated using Excel formulas. In addition, team member responses to the Kaizen event experience questions in the Kickoff Survey and Report Out Survey were used to calculate *Team Kaizen Experience*.

Upon conclusion of the study data collection window (October 2005 – August 2006), the researcher combined both spreadsheets into one master spreadsheet that contained all the survey data collected for the study period. As a final verification of the data, one Kickoff Survey and one Report Out Survey for each team were chosen at random and checked against the final, master spreadsheet (the "Individual-Level Data File"). In addition, the survey questions were checked for out of range responses – i.e., responses greater than 6 or less than 1 – using the Excel "max" and "min" functions. This master data file was imported into SPSS for further analyses described in the next sections.

In addition to the Kickoff Survey and Report Out Survey data, the Event Information Sheet data were imported into Excel for calculations including these variables. The data were transferred using "copy/paste" from the original, electronic Microsoft Word files of the Event Information Sheets. As a final verification, each Event Information Sheet was visually checked against the spreadsheet. Additional calculations were performed in this spreadsheet to calculate study variables from the raw Event Information Sheet data. These variables included:

Team Leader Experience; Work Area Routineness; Event Planning Process; Overall Perceived Success; Team Functional Heterogeneity; % of Goals Met; Tool Appropriateness and Tool Quality.

Finally, after all the data screening was complete, a separate master Excel spreadsheet was created that contained only the variables to be used in the team-level analysis (the "Team-Level Data File"). The data were transferred using "copy/paste" from the original, individual level spreadsheet. Once the "copy/paste" was completed, the variables were visually checked against the original spreadsheet for "copy/paste" errors.

Following the compilation of individual-level data into one central file, data were further screened to determine which survey responses and questions could be used in validation of the survey scales through the exploratory factor analysis. This screening process consisted of three major steps:

1. Calculating Kickoff Survey and Report Out Survey response rates for teams and determining whether any teams must be removed from the analysis due to low response rates – i.e., less than 50%.
2. Determining whether any of the individual surveys received show evidence of systematic response bias – i.e., survey fatigue – and should be removed from the analysis.
3. Determining whether any of the questions were problematic due to lack of variation in response or bimodality.

First, the response rate for each team was calculated. Teams with less than a 50% response rate on either the Kickoff Survey or the Report Out Survey were removed from the data set, since, especially given the small sizes of most teams – i.e., less than 10 persons – it was not clear that the responses received could be considered representative of the team. This resulted in the removal of four teams from the original 56 collected (see Tables 9 and 10 for further data on response rates), for an initial sample size of 52 teams. One additional team was removed later in the analysis, as will be described in later in this chapter. It should be noted that two of these four teams were excluded due to zero response on either or both the Kickoff Survey and Report Out Survey. For these teams, it appears that the facilitator either failed to administer the surveys or misplaced the completed surveys. The other two teams were excluded due to low response – i.e., between 0% and 50% – on the Report Out Survey. This data set included 347 individual Kickoff Surveys and 305 individual Report Out Surveys. As indicated in Appendix R, response rates were generally lower for the Report Out Survey than for the Kickoff Survey; this makes sense both due to the timing of administration and the relative length of the surveys.

Next, individual survey responses were analyzed to determine whether any respondents demonstrated clear evidence of systematic bias in response – i.e., survey fatigue. Survey fatigue occurs when a respondent becomes tired of answering survey questions and responds with the same value to a series of survey questions, without giving careful thought to each question. Survey fatigue is more common on longer surveys and could also be influenced by perceived lack of time to complete the survey. Thus, survey fatigue, if it occurred, was expected to be more problematic for the Report Out Survey (39 questions total) than the Kickoff Survey (19 questions total) both due to the number of questions and the timing of administration – i.e., the Kickoff Survey is administered during the beginning of the event, as part of the official event activities, while the Report Out Survey is administered at the report out meeting when participants may be tired from the event and eager to return to their normal jobs. Since judging survey fatigue is always somewhat subjective, a conservative criterion was applied – i.e., only "strong" cases of survey fatigue were discarded.

The same general procedure was used to screen for survey fatigue for both the Kickoff Survey and the Report Out Survey. For the Kickoff Survey, the standard deviation of response across all 17 scaled items in the survey was calculated for each respondent. Two respondents who demonstrated zero variation in response – i.e., answered all 17 of the Kickoff Survey questions with the same value – were removed from the data set. Next, the standard deviation across the last half (nine) of the scale questions was calculated. In general, any respondents who failed to demonstrate any variation in response across the last nine questions would have been removed from the data set. One additional respondent had a standard deviation of zero across the last nine questions. However, this respondent chose not to answer five of the final nine questions, thus suggesting that he/she was giving thought to his/her response, since the questions missing a response were not all immediately adjacent. This respondent, therefore, was retained in the data set. As a final verification, all respondents with a low standard deviation – i.e., less than 0.5 – were examined for suspicious response patterns. Although seven additional respondents (out of 347) had a standard deviation of less than 0.5, none of these cases demonstrated clear response bias. All seven respondents demonstrated some variation in response in the last nine questions and all but one demonstrated variation in the last two questions, one of which was reverse coded. Thus, a total of 345 Kickoff Surveys were retained for further analysis. A similar procedure was used for the Report Out Survey. In identifying response bias, particular attention was given to identifying respondents who failed to demonstrate variation on page two of the survey (i.e., questions 19-35), since this was a pattern the researcher had noted on a few surveys during data entry. Five respondents were removed due to zero variation – i.e., a standard deviation of zero. An additional 20 responses were removed due to evidence of systematic bias – i.e., no variation in response across questions 19-35 – for a total of 25 responses removed. As a final verification, all 46 additional respondents with standard deviations of less than 0.5 were visually checked, but no additional respondents were removed from the data set. Thus, a total of 280 Report Out Surveys were retained for further analysis.

Finally, each scale item in the Kickoff Survey and Report Out Survey was analyzed to determine whether any of the questions demonstrated a lack of variation or other statistical properties that would make it unsuitable for analyze. First, the mode of each survey question was examined to determine if any questions had a mode equal to the high scale value – i.e. 6. None of the survey questions had a mode equal to 6. However, the majority of questions did have a mode equal to 5 (11 out of 17 scaled Kickoff Survey questions and 34 out of 35 scaled Report Out Survey questions). Next, the mean, median, maximum, minimum and standard deviation were calculated for each question. In addition, histograms and tests of normality were conducted to determine distributional properties. For the Kickoff Survey, all questions had relatively symmetrical distributions – the mean, median and modes were relatively similar and the histograms demonstrated no evidence of bimodality or other abnormal response patterns. In addition, examination of the minimum and maximum values indicated that, in general, respondents were using the entire survey scale. The maximum response for all questions was 6, and the minimum for all but one question was 1. The other question had a minimum of 2. However, as has already been noted, the responses were negatively skewed – i.e., with the tail of the distribution pointing toward the low end of the survey scale; thus questions failed formal tests of normality. However, the relatively symmetry suggests that departures from normality are not extreme and the questions can be used in statistical analysis. For the Report Out Survey, 34 out of 35 questions had a mode of 5 and the other question had a mode of 4. Again, examination of the minimum and maximum values indicated that respondents were using the entire survey scale for most questions. The maximum response for all questions was 6; 25 of the questions had a minimum of 1, nine had a minimum of 2, and one had a minimum of 3. Again, question distributions appeared relatively symmetric, although negatively skewed; thus questions failed formal tests of normality.

3.5 Factor Analysis of Survey Scales

Following initial data screening, exploratory factor analysis was conducted to analyze the construct validity of the Kickoff Survey and Report Out Survey scales – i.e., the degree to which scale items appear to measure the same underlying characteristics, as indicated by a relatively high degree of correlation among items within a given scale compared to items that are part of other scales (Cronbach & Meehl, 1955; Martin, 1996; Leedy & Ormrod, 2005). Although a common step in most research involving survey questionnaires, factor analysis is particularly important for assessing the construct validity of survey scales that are new or have been modified in their original wording, as in the current research. Factor analysis is necessary for providing evidence that survey scales are, indeed, valid measures of unique variables of interest, and is a prerequisite to further analysis.

The Kickoff Survey and Report Out Survey scales were analyzed separately, due to differences in time of administration – i.e., the Kickoff Survey and Report Out Survey are administered at two separate times. In addition, the outcome scales in the Report Out Survey – i.e., *Overall Perceived Success*, KSAs and *Impact on Area* – were analyzed separately from the independent variable scales – i.e., *Team Autonomy*, *Management Support*, *Action Orientation* and *Internal Processes*. This was done since the outcome variables are theorized to be direct linear combinations of the independent variables (see Figure 1) and thus one might expect cross-loadings between outcomes variables and independent variables. In each factor analysis, a listwise exclusion method was used for missing data. This resulted in a somewhat lower n for each factor analysis than the total number of Kickoff Surveys or Report Out Surveys - $n_{fa} = 307$ for the Kickoff Survey factor analysis (original $n = 345$), $n_{fa} = 252$ for the Report Out Survey independent variable factor analysis (original $n = 280$) and $n_{fa} = 250$ for the Report Out Survey outcome variable factor analysis (original $n = 280$). References on factor analysis suggest a minimum ratio of observations versus items of 2:1 (Kline, 1994), with the preferred ratio closer to 10:1 (Tinsley & Tinsley, 1998; Kline, 1994). With the listwise exclusion, the observed ratio of observations versus items was approximately 22:1 for the Kickoff Survey analysis, approximately 15:1 for the Report Out Survey outcome variable analysis, and approximately 14:1 for the Report Out Survey independent variable analysis.

Principal components was used as the extraction method (Johnson, 1998). Although maximum likelihood can be preferred in terms of generalizability of results (Field, 2005), it is also less robust to departures from multivariate normality (Johnson, 1998; Fabrigar et al., 1999). As has already been described, univariate analyses of the Kickoff Survey and Report Out Survey questions suggested departures from univariate normality, thus making the assumption of multivariate normality untenable. Fortunately, most multivariate methods are quite robust to departures from multivariate normality unless the departures are severe; however, principal components was selected since it is more robust than maximum likelihood methods (Johnson, 1998). An oblique rotation method was used (oblimin with a delta of zero – i.e., direct quartimin rotation, see Field, 2005), since theory, as depicted in the preliminary research model (Figure 1), suggests that the underlying factors may be correlated (Fabringar et al., 1999, Finch, 2006). Because several of the hypotheses will be tested through regression, an orthogonal rotation could appear to offer an advantage in that it could result in less correlated factors. However, the research model structure assumes that independent variables may be correlated, thus an orthogonal rotation would not be consistent with the model assumptions. In addition, if variables are uncorrelated, oblique rotation methods and orthogonal rotations will result in similar results (Fabringar et al., 1999). Many other studies that have used regression analysis have also explicitly acknowledged the potential correlation between predictors measured through survey scales by performing oblique rotations during exploratory factor analysis to test scale construct validity (for some recent applications see Reed et al., 2006; Foley et al., 2006; Lawyer et al., 2006 and Breitkopf, 2006). However, in any study where correlations between predictors may exist, multicollinearity issues should be anticipated and addressed during the regression modeling process (Neter et al., 1996). As a final step in the factor analyses, the correlation matrix of the questions included in each analysis was examined to verify that the questions displayed some correlations and the determinant of the correlation matrix was relatively small – i.e., close to zero (Johnson, 1998), which are necessary for factor analysis to produce useful results.

It should be noted here that the presumed error structure of the data – i.e., individuals within teams within organizations – violates the underlying assumption of most multivariate analysis methods that the observations are independent (Johnson, 1998). This can be problematic because it can distort the standard error estimates used in statistical significance tests, and, depending on the nature of the violation and the relationship being tested, this can lead to spurious results of significance tests (Type I error) or, conversely, reduced power (Type II error) (Kenny & Judd, 1986). However, the current analysis is descriptive, rather than inferential. The goal is to analyze patterns of correlations and to determine which groups of questions appear to be more highly related to one another for the purpose of validating the survey scales – i.e., determining which questions appear related to the same underlying construct and can be average together. It is not intended to be used to test for the statistical significance of factor loadings and actual factor loadings are not used in calculating scale values. Instead the average of the items in the scale is used. Since we cannot soundly test for statistical significance, only questions with factor loadings that are relatively high in magnitude – i.e., greater than 0.500 – should be considered, versus a lower cutoff point, such as a 0.300. In addition, questions with cross-loadings greater than 0.300 were removed from the analysis (Kline, 1994).

Use of factor analysis as an exploratory tool of this type is appropriate, as long as formal significance tests of factor loadings are not conducted (Schulman, 2006, personal communication). Other approaches to dealing with the nested error structure include conducting the factor analysis at a higher level (e.g., see Van den Bree et al., 1996 where dyad averages were used to analyze data from sets of twins) or the exclusion of related observations (e.g., Hammer et al., 2000). For instance, in the current analysis, one observation from each team could have been randomly selected to include in the factor analysis. However, either of these options would have drastically reduced the sample size – i.e., to 52 teams.

The established heuristic of extracting all factors with eigenvalues greater than 1.0 (Johnson, 1998; Field, 2005) was used to determine the number of factors. However, in the case of the Report Out Survey outcome variable analysis, the third factor had an eigenvalue equal to 0.999 – due to the closeness to 1.0, this was rounded up to 1.0 and a third factor was included. In addition, scree plots were examined to determine if there was a clear break in factor importance, and the magnitude of the next highest eigenvalue not extracted was examined to determine how close this value was to one. In cases where the next highest eigenvalue not extracted was close to 1.0 – i.e., greater than 0.9 – it might make sense to extract an additional factor, both to test the robustness of the factor solution and to potentially explain any loadings in the factor solution that were difficult to interpret. In the Kickoff Survey factor analysis and Report Out Survey Outcome Variable factor analysis, the next highest eigenvalue not extracted was around 0.7. Therefore, no additional factors were examined after the initial solution. However, for the Report Out Survey Independent Variable factor analysis, the next highest eigenvalue not extracted was greater than 0.9. Therefore, following the initial solution, the next highest eigenvalue was extracted to test the robustness of the initial solution. Tables 11, 12 and 14 present the resulting pattern matrices from the factor analyses, as executed in SPSS. To improve the readability of the tables, factor loadings of less than 0.100 have been suppressed, following common practice.

3.5.1 Factor Analysis of Kickoff Survey Scales

As shown in Table 11, all of the questions in the Kickoff Survey loaded as theorized, providing support for the construct validity of the Kickoff Survey scales. Meaningful loadings – i.e., greater than 0.500, where all cross-loadings are less than 0.300 – are shown in bold. All questions load highly onto the intended scales – i.e., the minimum observed loading was 0.711 – and have low loadings on the other two scales – i.e., the maximum observed cross-loading was 0.154. Thus, all of the original items in each of the *Goal Clarity* (GC), *Goal Difficulty* (GDF) and Commitment to Goals (ACC) scales were retained for further analysis.

Table 11. Pattern Matrix for Factor Analysis of Kickoff Survey Scales[1]

	Component		
	1	2	3
ACC1	**.774**	.128	
ACC6	**.770**		
ACC5	**.756**		
ACC4	**.722**		-.154
ACC3	**.715**		
ACC2	**.711**		-.130
GDF2		**.862**	
GDF1		**.843**	
GDF3		**.747**	
GDF4	-.122	**.723**	
GC3			**-.881**
GC1			**-.875**
GC4			**-.808**
GC2			**-.754**

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.
Rotation converged in 5 iterations.

3.5.2 Factor Analysis of Report Out Survey Scales – Independent Variables

Table 12 shows the loadings of the independent variables measured in the Report Out Surveys – i.e., *Team Autonomy*, *Management Support*, *Action Orientation* and *Internal Processes*. Meaningful loadings – i.e., greater than 0.500, where all cross-loadings are less than 0.300 – are shown in bold. In general, the factor loadings support the construct validity of the four survey scales; however, the factor loadings did result in the removal of some items that did not load highly enough onto the intended scale or displayed high cross-loadings on other scales. Although most items loaded highly on one scale only, some items displayed primary loadings less than 0.500 and/or cross-loadings greater than 0.300.

All five *Internal Processes* (IP) items loaded highly onto a single factor (minimum observed loading = 0.591; maximum observed cross loading was 0.201). Three out of five of the *Management Support* (MS) items loaded together (minimum observed loading = 0.655; maximum observed cross-loading = 0.148). These three items –

[1] Negative cross-loadings are due to the orientation of the axes from the factor rotation. They do not (necessarily) indicate that the question has a negative relationship to the construct of interest -- particularly in cases such as the GC scale where all the high loadings were negative.

MS2, MS3 and MS5 – relate to sufficiency of materials and supplies, equipment and help from others in the organization. MS4, which is related to sufficiency of help from the facilitator, nearly loaded onto the *Team Autonomy* (TA) factor (0.481), while MS1, which relates to sufficiency contact with senior management did not load meaningfully onto a single factor. Instead, MS1 displayed high cross-loadings – i.e., greater than 0.300 – on both MS and TA. This result is logical; however, MS1 should be excluded from further analysis because it does not load clearly onto one scale.

Only two of the *Action Orientation* (AO) items loaded together (minimum observed loading = 0.733; maximum observed cross-loading = 0.223). This result makes sense, because these two items (AO1 and AO2) were the most similar in focus of the items in the AO scale: AO1 = "Our Kaizen event team spent as much time as possible in the work area" and AO2 = "Our Kaizen event team spent very little time in our meeting room." Of the remaining items, AO4 ("Our Kaizen event team spent a lot of time discussing ideas before trying them out in the work area") nearly loaded onto the IP factor (0.423), which makes sense since this item relates to team processes – i.e., discussion. However, due to the relatively low loading, AO4 will be excluded from further analysis. AO3 ("Our Kaizen event team tried out changes to the work area right after we thought of them") displayed a high loading on the TA factor (0.763) but also displayed a high cross-loading on the AO factor (0.301). Again, this result makes sense, since this item refers both to the act of making changes immediately versus waiting (AO) and the ability to do so (TA).

Finally, three of the *Team Autonomy* (TA) items loaded together (TA1, TA2 and TA3). The fourth TA item (TA4) did not load highly enough to be included (0.479). In addition, as previously mentioned, AO3 had a high loading on the TA scale (0.763), but also displayed a high cross-loading on the AO scale (0.301), as well as a moderate cross-loading on the MS (0.235). Although the magnitude of the primary loading on TA was high, the relatively high cross-loadings indicate that the item does not load cleanly onto a single scale – i.e., it displays relatively strong links to two additional constructs. Thus, despite the high primary loading on the TA scale, a conservative approach is taken and AO3 is excluded from further analysis. Finally, as mentioned, MS4 nearly loaded onto the TA factor (0.481). The most logical explanation for the near loading of MS4, which relates to the sufficiency of help from the facilitator, onto the TA factor appears to be that the facilitation process is tightly linked to enabling team autonomy.

The revised survey scales, based on the factor analysis results, are presented in Table 12.

Table 12. Pattern Matrix for Factor Analysis of Report Out Survey Scales – Independent Variables

	Component			
	1	2	3	4
IP3	**.882**		.152	
IP2	**.872**			
IP4	**.850**			.109
IP5	**.617**		-.102	
IP1	**.591**	-.158		-.201
AO4	-.423		.119	
AO2		**.861**		
AO1	.223	**.733**	-.169	
MS3			**-.891**	
MS2			**-.839**	
MS5	.115	.148	**-.655**	-.121
MS1	.160		-.396	-.350
TA3				**-.830**
TA2	.132	-.139		**-.794**
AO3		.301	.235	-.763
TA1			-.276	**-.508**
MS4	.153	-.181	-.280	-.481
TA4			-.266	-.479

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.
Rotation converged in 13 iterations.

Table 13. Revised Report Out Survey Scales – Independent Variables

Scale	Revised Item List
Internal Processes	• IP1: "Our team communicated openly." • IP2: "Our team valued each member's unique contributions." • IP3: "Our team respected each others' opinions." • IP4: "Our team respected each others' feelings." • IP5: "Our team valued the diversity in our team members."
Action Orientation	• AO1: "Our team spent as much time as possible in the work area." • AO2: "Our team spent very little time in our meeting room."
Management Support	• MS2: "Our team had enough materials and supplies to get our work done." • MS3: "Our team had enough equipment to get our work done." • MS5: "Our team had enough help from others in our organization to get our work done."
Team Autonomy	• TA1: "Our team had a lot of freedom in determining what changes to make to this work area." • TA2: "Our team had a lot of freedom in determining how to improve this work area." • TA3: "Our team was free to make changes to the work area as soon as we thought of them."

As mentioned, since the next highest eigenvalue not extracted for this factor analysis was relatively close to 1.0 (i.e., 0.914), a five factor solution was examined as a follow-up to the initial, four factor solution. The results support the robustness of the initial, four factor solution. The fifth factor that emerged was a trivial factor, consisting only of

AO4. The other items loaded as in the four factor solution, except for IP1, where the loading fell below 0.5 due to cross-loading with the AO4 factor (both items relate to team discussion), and TA1, which displayed a cross-loading of 0.301 on the MS factor. However, IP1 and TA1 were retained due to conceptual relationships to the rest of their scale, high scale reliability and lack of support for the five factor solution versus the four factor solution – i.e., the trivial fifth factor.

3.5.3 Factor Analysis of Report Out Survey Scales – Outcome Variables

Table 14 shows the loadings of the outcome variables measured in the Report Out Surveys – i.e., *Overall Perceived Success*, KSAs and *Impact on Area*. Meaningful loadings – i.e., greater than 0.500, where all cross-loadings are less than 0.300 – are shown in bold. In general, the construct validity of the survey scales was supported, with the exception of the *Understanding of CI* and *Skills* scales, which were not found to be empirically distinct. In addition, although most items loaded highly on one scale only, some individual items displayed primary loadings less than 0.500 and/or cross-loadings greater than 0.300, and thus were excluded from further analyses.

Three out of four *Impact on Area* (IMA) items loaded together (minimum observed loading = 0.794; maximum observed cross-loading = 0.213) – IMA1, IMA2 and IMA3. The fourth IMA item (IMA4) loaded onto a separate factor. This result makes sense due to the linguistic construction of IMA4, which will be discussed shortly.

In addition, two items from the *Attitude* (AT) scale loaded highly onto one factor – AT4, AT2 (minimum observed loading = 0.669; maximum observed cross-loading = 0.246), along with one item from the SK scale (SK4). This supports team member attitudes (AT) as a unique dimension of event impact. However, although the *Attitude* scale was originally designed as an overall measure of impact on two types of attitudes – affect for Kaizen events and employee attitudes toward work – the factor analysis suggests that the two types of attitudes are distinct and should be considered separately in the analyses. The two items loading highly onto the AT factor – AT2 and AT4 – directly describe enjoyment of Kaizen event activities. In addition, SK4 similarly describes a variable directly related to liking for Kaizen activities – the degree to which Kaizen event team members are comfortable working with others to identify improvements. However, the other AT items (AT1 and AT3), which loaded onto a different factor, describe the impact of the Kaizen event on participating employees' general attitudes toward work – i.e., task motivation.

Finally, the Knowledge of Continuous Improvement (UCI) and *Skills* (SK) items did not load onto separate factor, but instead most of these items loaded onto a single factor, which appears to be measuring employee gains in task-related – i.e., Kaizen process-related – KSAs (TKSA). While “knowledge” and “skills” are conceptually distinct in the literature, it appears that in practice they are very highly related. Both refer more to “technical” aspects of problem-solving capabilities, as distinct from affective attitudinal response toward Kaizen events. Thus, it is not surprising that these items loaded together. The only exceptions were SK2 and SK4. SK4 loaded onto the AT factor, as previously discussed. SK2 displayed high cross-loadings – i.e., greater than 0.400 – on both the TKSA factor and the IMA factor. This result makes sense since this particular item seeks to ascertain the extent to which team members have the ability to measure the impact of changes to the target work area – thus, this item addresses both an ability and perceived technical impact (IMA). In addition, AT1 and AT3, which relate to employees’ work – i.e., task – motivation also loaded onto the TKSA factor. Finally, IMA4 also loaded onto the TKSA factor. IMA4 says: “Overall, this Kaizen event helped people in this area work together to improve performance.” Thus, it appears that this item may actually be measuring employee gains in the ability to work in problem-solving – i.e., Kaizen event – teams, as specific TKSA.

Interestingly, the overall perceived success item (OVER) did not load cleanly onto the “technical success” measure IMA, but instead displayed a high cross-loading on the AT factor. This result makes sense and adds support to the proposition that the overall success of a Kaizen event is dependent on both technical and social system impact.

The revised survey scales, based on the factor analysis results, are presented in Table 15.

Table 14. Pattern Matrix for Factor Analysis of Report Out Survey Scales – Outcome Variables

	Component		
	1	2	3
UCI1	**.835**		
SK1	**.804**		
SK3	**.786**	-.204	
UCI4	**.764**	.101	
AT1	**.757**		
UCI3	**.757**		
UCI2	**.735**		.156
AT3	**.731**	.144	
IMA4	**.599**	.281	
SK2	.491	.440	-.224
IMA3		**.825**	.138
IMA2	.159	**.807**	-.156
IMA1		**.794**	.213
OVER		.659	.374
AT4			**.779**
AT2	.246		**.669**
SK4	.174	.195	**.536**

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.
Rotation converged in 6 iterations.

Table 15. Revised Report Out Survey Scales – Outcome Variables

Scale	Revised Item List
Task KSA	• UCI1: “Overall, this Kaizen event increased our team members' knowledge of what continuous improvement is.” • UC2: “In general, this Kaizen event increased our team members' knowledge of how continuous improvement can be applied.” • UCI3: “Overall, this Kaizen event increased our team members' knowledge of the need for continuous improvement.” • UCI4: “In general, this Kaizen event increased our team members' knowledge of our role in continuous improvement.” • SK1: “Most of our team members can communicate new ideas about improvements as a result of participation in this Kaizen event.” • SK3: “Most of our team members gained new skills as a result of participation in this Kaizen event.” • AT1: “In general, this Kaizen event motivated the members of our team to perform better.” • AT3: “Overall, this Kaizen event increased our team members' interest in our work.” • IMA4: “Overall, this Kaizen event helped people in this area work together to improve performance.”
Impact on Area	• IMA1: “This Kaizen event had a positive effect on this work area.” • IMA2: “This work area improved measurably as a result of this Kaizen event.” • IMA3: “This Kaizen event has improved the performance of this work area.”
Attitude	• AT2: “Most of our team members liked being part of this Kaizen event.” • AT4: “Most members of our team would like to be part of Kaizen events in the future.” • SK4: “In general, our Kaizen event team members are comfortable working with others to identify improvements in this work area.”

3.6 Reliability of Revised Scales

Cronbach's alpha was calculated to assess the reliability of the revised survey scales. Cronbach's alpha is a measure of internal consistency – i.e., the degree to which items in a given scale are correlated (Cronbach, 1951). Only individuals who responded to all the items in a given scale were including in calculating Cronbach's alpha for that scale. Table 16 presents the resulting values. In addition to the calculation of Cronbach's alpha, the researcher also conducted analyses to determine the effects of deleting individual items on scale reliability. Although, in a few cases, very slightly apparent improvement resulted, none of these changes were deemed large enough to justify eliminating items. Except for the *Action Orientation*, all scales had values well above the suggested threshold of 0.700 (Nunnally, 1978). Seven of the 10 scales had values of 0.800 or greater and one scale had a value greater than 0.900. The *Action Orientation* scale was somewhat problematic since it contained only two items and had an alpha value of 0.640. Although this is less than the desired threshold, some sources suggest that a slightly lower alpha threshold – i.e., 0.600 – can be used for newly developed scales, such as the *Action Orientation* scale (DeVellis, 1991). Thus, the decision was made to retain *Action Orientation* in the present analysis, but to seek to develop additional *Action Orientation* scale items for future research.

Table 16. Cronbach's Alpha Values for Revised Survey Scales

Scale	Cronbach's Alpha	Largest Increase if Item Deleted
Goal Clarity	.860	.871 (GC2)
Goal Difficulty	.810	.790 (GDF4)
Affective Commitment to Change	.859	.858 (ACC3)
Internal Processes	.856	.846 (IP5)
Action Orientation	.640	n/a
Management Support	.802	.807 (MS5)
Team Autonomy	.791	.732 (TA3)
Task KSA	.929	.930 (SK3)
Impact on Area	.866	.857 (IMA2)
Attitude	.738	.721 (SK4)

Following the calculation of Cronbach's alpha, scale averages for each individual in the data set were calculated using the revised scales. In addition, at this point the decision was made to substitute the averages of the other items in the scale for missing values in cases where individuals were missing only one item in a given scale. This allowed the calculation of a scale value for that individual, with minimum risk to data integrity, given the high inter-item correlation – i.e., Cronbach's alpha values – between items in the scale. This within-person replacement approach – also called a "person-mean" approach (Roth & Switzer, 1999) – has been demonstrated to be superior to other approaches for replacing missing data in terms of minimizing bias while maintaining power (Roth et al., 1999). However, it is important to note that, in general, most approaches to replacing item-level missing data – e.g., replacement with the item mean across all cases, regression from the other items in the scale, replacement with the within-person mean of other items in the scale, etc. – seem to only introduce a small amount of bias, unless a large percentage of the data is missing (Roth el al., 1999). Conversely, considerable power could be lost if pairwise or listwise deletion were practiced in cases where only single-items within a scale are missing. In cases where individuals skipped multiple questions in a given scale, the scale average was not calculated, since the risk to data integrity was deemed too great. These individual-level scale averages were used for additional exploratory analyses in preparation for hypothesis testing.

At this point, one additional team – Team 101 from Company E – was dropped from the analysis due to the fact that there was only one scale average for each of the three Kickoff Survey scales. Only one individual fully completed the Kickoff Survey while others skipped multiple questions. This problem was noted during initial screening; however, this team was retained until the completion of the factor analysis and the calculation of Cronbach's alpha, because it was not absolutely clear earlier what final set of variables and questions would be retained – i.e., the Kickoff Survey scales could have been significantly modified or even potentially removed if scale reliability were very low.

3.7 Aggregation of Survey Data to Team-Level

Following the analysis of the construct validity of the survey scales, calculation of the reliability of the revised scales, and the calculation of new individual-level scale averages, the next step in preparing the data for modeling was analyzing whether the data could justifiably be aggregated to the team level. The intended unit of analysis in this research is the team. However, since the data were collected on perceptual measures, it was necessary to collect data at the individual-level – i.e., individuals within teams. Measures were designed to reflect group-level attributes, thus the group was used as the referent in all survey questions. This is a called a "referred shift" composition model (Chan, 1998). If the measures work as designed – i.e., are really measuring shared group properties – there should be more variation across teams than within teams – i.e., the one-way ANOVA with team as a main effect should be significant – and there should be a relatively high degree of consensus within the team, measured by analyzing the extent of interrater agreement. If these criteria are met, this suggests that individual level data can be averaged such that the group mean score reflects the group value for the variable of interest. If the variable of interest is truly shared at a group level, the group mean will represent a more reliable measure than the individual values (Bliese, 2000).

Thus, in the current research, it was necessary to empirically demonstrate support for aggregation through statistical analysis of the data collected. As described by Bliese (2000), measures used to evaluate support for aggregation fall into two basic types: 1) measures that compare observed within-group variance to a theoretical random variance – i.e., agreement indices; and 2) measures that compare observed within-group variance to observed between-group variance – i.e., reliability/non-independence indices.[2]

The most commonly used measure of within-group agreement is r_{wg} – or more properly, $r_{wg(j)}$, where j is the number of items in the scale (James et al., 1993, 1984) . However, the general term r_{wg} is more commonly used and will be used be used throughout this work to refer to the multi-item within-group agreement index for a given scale. The r_{wg} statistic has been widely used (Van Mierlo et al., 2006), including recent uses in team applications as the sole justification for aggregation (e.g., see Passos and Caetano, 2003). The r_{wg} statistic measures the consistency of ratings within a group – i.e., team – by comparing the within-group variance (s_x^2) to a theoretical expected random variance (σ_E^2). Generally, the uniform distribution is used to estimate σ_E^2, since this reflects the expected response distribution for an interval scale with absolutely no response bias (e.g., no social desirability effects, central tendency, etc.). However, other distributions may be used for σ_E^2 to account for the possibility of response bias (James et al., 1984). Values for r_{wg} range from zero to one, reflecting the extent to which s_x^2 differs from σ_E^2. The maximum r_{wg} value of 1.0 is achieved when s_x^2 is zero – i.e., there is perfect within-group agreement. The minimum value occurs when s_x^2 equals or exceeds σ_E^2. James Demaree, and Wolf (1993) and (1984) prescribe that when s_x^2 exceeds σ_E^2, r_{wg} be set equal to zero for interpretation purposes.[3] Thus, when the within-group variance (s_x^2) is much smaller than σ_E^2 a relatively large r_{wg} value occurs. Generally, 0.7 is used as a threshold value to justify aggregation (George, 1990; Klein & Kozlowski, 2000; Van Mierlo et al., 2006). Although statistical tests of the

[2] Reliability and non-independence indices use the same formulations – e.g.., ICC(1), ICC(2) – and thus measure the same attributes. However, the term "reliability" is often used when a dependent variable is being analyzed, while the term "non-independence" is often used when an independent variable is being analyzed (Bliese, 2000). Throughout the rest of this work, the term "reliability" is used to refer to these types of indices.

[3] The actual r_{wg} values in this case can be negative or, conversely, greater than one, if the observed variance is much greater than the theoretical variance such that there are negative terms in both the numerator and denominator of the r_{wg} calculation – see Equation 3

significance of r_{wg} can be performed, in the current research these tests would be subject to errors from the presumed error structure of the data – i.e., teams within organizations.

One limitation of r_{wg} as used in some studies is that, in many cases, the uniform distribution may not be the best choice for the theoretical expected random distribution, due to phenomena such as central tendency in response or social desirability effects, which may artificially restrict the range of responses, resulting in overly liberal values of r_{wg}. Thus, the values achieved for r_{wg} using the uniform distribution provide a practical "upper bound" on r_{wg} (Bliese, 2000). In addition to calculating r_{wg} using the uniform distribution, the researcher can estimate a more conservative practical "lower bound" value for r_{wg} by selecting a second theoretical expected distribution that accounts for the likely type of response bias. For questions where a social desirability bias is likely, a skewed distribution can be used to estimate r_{wg}. For questions where a central tendency bias is likely, a triangular distribution can be used to estimate r_{wg} (James et al., 1984). Another conservative technique for estimating r_{wg} is random group resampling (RGR) (Bliese et al., 1994), an estimation technique based on bootstrapping, in which within-group variance of actual groups is compared to the within-group variance of randomly formed pseudo-groups. This technique was not used in the present research due to concerns about potential bias in the pseudo-group variance estimates due to the nested structure of the data – i.e., teams within organizations. Finally, although r_{wg} can be useful as an index of within-group agreement, r_{wg} does not demonstrate whether the variable of interest varies across groups (Bliese, 2000). Although nothing conceptually requires group-level measures to vary across the groups in the sample in order to be "group-level," variables that demonstrate zero variance across groups are not useful in analysis. This is why the calculation of reliability measures can be especially helpful in evaluating study variables.

Another, relatively new, measure of interrater agreement is the average deviation index (AD) (Burke et al., 1999; Burke & Dunlap, 2002; Dunlap et al., 2003). More precisely, AD measures the extent of disagreement among raters – i.e., the average deviation from the mean response among a group of raters. Although AD has some conceptual differences from r_{wg} – specifically, AD is expressed in the same units as the underlying scale, while r_{wg} is unitless, ranging from 0 to 1 – AD has been found to be highly correlated with r_{wg} ($r = -0.91$) (Burke et al., 1999). In addition, similar to r_{wg}, the interpretation of AD is based on comparing the observed value of AD to the expected value of AD under a uniform response distribution (Burke & Dunlap, 2002). However, AD is currently less flexible than r_{wg} in modeling other expected response distributions. No formulas or heuristic thresholds appear to have been developed for comparing the observed AD to its expected value under other response distributions. Thus, in the current research, r_{wg} will be used instead of AD since they provide similar information and, in addition, r_{wg} will allow the substitution of additional expected response distributions for the uniform distribution in order to examine the practical "lower bound" of interrater agreement.

One reliability measure used to evaluate support for aggregation is the intraclass correlation coefficient (1), or ICC(1). ICC(1) is a measure of the amount of lower-level variance that can be explained by group membership (Bliese, 2000). There are actually two ICC(1) measures in use in organizational research, both of which are, confusingly, referred to only as ICC(1). The first, is the classic ICC(1) measure that has been historically used in statistics and social sciences research, particularly in applied psychology (e.g., Bartko, 1976; Kenny & La Voie, 1985). This measure is calculated from the mean squares resulting from ANOVA and ranges from -1 to 1.[4] In this formulation, ICC(1) values are negative when within-group variance exceeds between group variance, such that ICC(1) is really a measure of the difference between the between-group mean square and within-group mean square from ANOVA. The second ICC(1) measure is calculated from the variance components resulting from an unconditional – i.e., means only – random effects model[5], and is commonly used in conjunction with hierarchical linear modeling (HLM) (Raudenbush & Byrk, 2002). This ICC(1) is a more direct calculation of the proportion of total level-specific variation – i.e., between-group variation plus within-group variation for the given level of analysis (Raudenbush & Byrk, 2002) – that can be accounted for by group membership. This is similar to the interpretation of the eta-squared measure in the one-way ANOVA, which will be discussed more presently. Although both formulations essentially measure the same attribute, they use different scales. The variance component formulation ranges from 0 to 1 only. In the variance components formulation, a value of zero is only attained when there is no between group variance. The ICC(1) value will still be positive even when total within-group variance exceeds between-group variance.

In both formulations, larger ICC(1) values indicate greater homogeneity within groups on the measure of interest. Although there are no clear limits for when ICC(1) is large enough to justify aggregation (Schneider et al., 1998), as a rule of thumb, values greater than 0.1 are often taken to indicate justification for aggregation (James,

[4] The range is actually -1 to 1 for dyads and -1/(k-1) to 1 for groups larger than dyads, where k is the average group size (Kenny & Judd, 1986)

[5] In the variance component formulation ICC(1) = $\frac{\sigma_B^2}{\sigma_B^2 + \sigma_W^2}$ where σ_B^2 is the variance between groups and σ_w^2 is the variance within groups. The mean square formulation of ICC(1) is shown as Equation 1.

1982; Schneider et al., 1998; Molleman, 2005). Although this rule of thumb is applied equally to both formulations, it obviously reflects a more stringent test of association for the means squares formulation of ICC(1). An ICC(1) of 0.20 has been suggested as demonstrating a strong group-level association (Molleman, 2005), and again, this rule of thumb is applied to both formulations, even though it reflects a different strength of association for each measure. Another method used to justify aggregation is by testing the significance of ICC(1), which can be used to test either formulation. In the mean square formulation of ICC(1), a statistically significant ANOVA with group as the main effect indicates that ICC(1) is significant (Bliese, 2000; Klein and Kozlowski, 2000), providing justification for aggregation. Testing the significance of ICC(1) is widely used in organizational research to justify aggregation (e.g., Van Mierlo et al., 2006; Sarin and McDermott, 2003).

Another technique used to evaluate support for aggregation is within-and-between analysis I (WABA I) (Dansereau et al., 1984). In addition to a statistical significance tests using ANOVA, a formal WABA I approach also uses a "practical significance test," which is based on the between group eta-squared value from one-way ANOVA[6]. Similar to ICC(1), the eta-squared value used in the practical significance test is a measure of the amount of lower-level variance that can be explained by group membership (Klein & Kozlowski, 2000). In fact, when group sizes are large (i.e., larger than 25), eta-squared values are approximately equal to ICC(1) values calculated using the variance component formulation. However, unlike ICC(1), eta-squared is affected by group size, and when group sizes are small, eta-squared overestimates between-group variance, resulting in inflated (i.e., biased) eta-squared values (Bliese & Halverson, 1998). Although Bliese and Halverson (1998) show that eta-squared values can be substantially biased for groups as large as 25, they are strongly upwardly biased for groups of 10 or less. Because all of the Kaizen events in this research had team sizes smaller than 25 and most were even smaller than 10, a formal WABA approach using direct examination of eta-squared values and/or practical significance tests was not used.

---

[6] Specifically, the practical significance test is based on the expected ratio of the between-group eta-correlation, which is the square root of eta-squared, to the within-group eta-correlation (Dansereau et al., 1984) under different theoretical relationships – i.e., wholes, parts, equivocal, inexplicable/null (see Dansereau et al., 1984 for more details on each of these conditions). η_B^2= SSB/SSTO, while η_w^2 = SSW/SSTO = SSE/SSTO. Thus, the eta-squared values are interdependent in that they sum to 1.0, if only one main effect – group – is being examined at a time. Through both mathematical substitution and simulation, Bliese and Halverson (1998) demonstrate that the expected between-group eta-squared values under the conditions of a known amount of between group variance – i.e., a known group effect – differ as a function of average group size. Thus, the expected ratio of the between eta-correlation and within eta-correlation, given a certain type of relationship, also differs as function of group size, and the suggested practical significance "confidence intervals" (Dansereau et al., 1984) for determining what type of relationship exists are only useful for dyads (see Bliese & Halverson, 1998 for more details). In addition, as described more above, direct examination of η_B^2 can also be misleading due to this group size effect.

This research used both the classic, mean square formulation of ICC(1) and r_{wg} as justification for aggregation. The primary intent here is to analyze whether, as designed, the measures do appear to result in less variation within versus across groups and can be aggregated for further analyses to test study hypotheses. Thus, it appears that the mean square formulation is better for these purposes, because it more directly contrasts the between-group and within-group mean squares. For each survey variable, the absolute magnitude of ICC(1) was calculated and the significance tests from ANOVA was examined to evaluate the extent to which there appears to be greater variation within versus across groups. As previously described, ICC(1) is essentially a measure of the proportion of individual-level variation that can be explained by group membership. In essence, this amounts to a measure of the extent to which individuals belonging to the same team are more similar to one another in their ratings than to individuals in others teams (Kenny & La Voie, 1985). Typically, if there is only one level of nesting – e.g., individuals within teams – a one-way ANOVA with group as a main effect can be used both to calculate ICC(1) and to test its significance. As Equation 1 shows, ICC(1) is calculated using the group effect mean square and error – i.e., within group – mean square from ANOVA output. However, in the current research there are two levels of nesting – individuals within groups and groups within organizations. Therefore, a traditional one-way ANOVA with group as the main effect cannot be used, since this will result in inaccurate mean square estimates, if, as suspected, there is correlation between groups within organizations (Kenny & Judd, 1986). To control for organizational-level effects and to correctly partition the variance – i.e., correctly calculate mean squares – it was necessary to use a nested ANOVA where both *organization* effects and *group (organization)* effects were included in the model (e.g. DeCoster, 2002). Although the approach of using a nested ANOVA to calculate the mean square derivation of ICC(1) does not appear to have been used much to date in team research, nested ANOVA has been used to examine group effects in the presence of multiple levels of nesting in other fields, such as experimental psychology (e.g., DeCoster, 2002; Jeten et al., 2002), biology (e.g., Poulsen, 2002) and physical anthropology (see Smith, 1994 for an overview of the approach and citations of other studies).

The model was executed using the SPSS "Mixed Models" procedure in SPSS 11.0. Table 17 presents the p value for the nested ANOVA and the ICC(1) values as calculated using the ANOVA output in the formula from Bartko (1976) (Equation 1), where k is the average group size and MSB and MSW are the outputs from the ANOVA with group – i.e., team – nested within organization. For rigor, two approaches presented in the literature were used for estimating k in the ICC(1) calculations. The first and simplest approach is to use the raw average group size –

i.e., N/J, where J is the number of teams (i.e., 51) and N is the total number of individuals. However, using this k may slightly bias the results when group sizes are very unequal – i.e., the corrected k is very different from the raw average group size. Although this did not appear to be the case in the current analysis, ICC(1) was also calculated using the formula for a corrected k reported by Blalock (1972) and Haggard (1958) (see Equation 2), where is the N_i number of individuals in each team. As can be seen, the ICC(1) values using the corrected k were slightly greater than the original ICC(1) values, but only varied slightly from the original values – i.e., in the third or fourth decimal place – since the corrected k was not very different from the raw average group size.

$$ICC(1) = \frac{MSB - MSW}{MSB + [(k-1) * MSW]} \quad \textbf{(1)}$$

$$k = \frac{1}{J-1}\left(\sum_{i=1}^{J} N_i - \frac{\sum_{i=1}^{J} N_i^2}{\sum_{i=1}^{J} N_i} \right) \quad \textbf{(2)}$$

Table 17. Nested ANOVA p-values and ICC(1) Values for Survey Scales

Scale	***ANOVA p***	***N***	***k average***	***k corrected***	***ICC(1) using k average***	***ICC(1) using k corrected***
Goal Clarity	.0465	336	6.59	6.56	.060	.060
Goal Difficulty	.0000	338	6.63	6.60	.216	.217
Affective Commitment to Change	.0000	338	6.63	6.60	.173	.174
Internal Processes	.0000	276	5.41	5.39	.326	.327
Action Orientation	.0000	276	5.41	5.39	.373	.374
Management Support	.0030	276	5.41	5.39	.125	.126
Team Autonomy	.0000	274	5.37	5.35	.185	.186
Task KSA	.0000	274	5.37	5.35	.121	.121
Impact on Area	.0000	268	5.25	5.23	.429	.430
Attitude	.0000	276	5.41	5.35	.300	.303

As can be seen from Table 17, all 10 ANOVAs were significant, indicating that all ICC(1) values were statistically significant. However, the ICC(1) for *Goal Clarity* was barely significant (p = 0.0465). Furthermore, if a Bonferroni correction is used to adjust the alpha value for the number of tests (i.e., 10 tests, corrected alpha = 0.05/10 = 0.005 or even 0.10/10 = 0.01), the ICC(1) for *Goal Clarity* would not be considered significant, although all other ICC(1) values would be considered significant. This indicates that the justification for aggregation for *Goal Clarity* is weaker than for the other nine variables in terms of the ICC(1) values – i.e., the contrast of between-group and within-group mean squares. In addition, except for *Goal Clarity*, all ICC(1) values were greater than the recommended rule of thumb value of 0.1. In addition, five ICC(1) values were greater than 0.2, which is a rule of thumb value indicating strong intraclass correlation. However, in addition to having a relatively high p-value, the ICC(1) value for *Goal Clarity* was somewhat lower than 0.1 – i.e., 0.060.

In addition to calculating ICC(1), the r_{wg} for each team was calculated for each survey scale. Two approaches were used to calculate r_{wg}, using the basic formula from James, Demaree and Wolf (1984) (Equation 3). First, the uniform distribution was used for σ_E^2 in order to estimate r_{wg} for situations with no response bias. As mentioned, this provides a practical "upper bound" on r_{wg}. Next, to estimate a practical "lower bound" on r_{wg}, a second distribution was substituted for σ_E^2 to reflect the likely direction of any response bias. For most of the survey constructs, one might posit a moderate skew due to social desirability – i.e., more responses on the positive – "agree" – side of the survey scale. This distribution is chosen based both on theoretical reasoning, as well as empirical observation of the range of data collected – i.e., initial screening of the data revealed that respondents were using the full six point scale for most questions, although the observed question distributions were negatively skewed. However, for two scales, *Goal Difficulty* and *Action Orientation*, it seems like a central tendency bias may be more likely, since there is no reason to suspect social desirability bias on these scales. Thus, a moderately skewed distribution was used for *Goal Clarity*, *Affective Commitment to Change*, *Internal Processes*, *Management Support*, *Team Autonomy*, *Task KSA*, *Impact on Area*, and *Attitude*, and a triangular distribution was used for *Goal Difficulty* and *Action Orientation*.

$$r_{wg} = \frac{J[1-(\overline{s_{xj}^{\,2}}/\sigma_E^{\,2})]}{J[1-(\overline{s_{xj}^{\,2}}/\sigma_E^{\,2})]+(\overline{s_{xj}^{\,2}}/\sigma_E^{\,2})} \qquad \textbf{(3)}$$

Where J = the number of items in scale and s_{xj}^2 = the average of the within-team variances for each of the J items in the scale.

For the uniform distribution, $\sigma_E^2 = \sigma_{EU}^2 = (A^2 - 1)/12$, which is the expected variance of the uniform distribution on a scale with A response intervals. Since all the scales in this research have six response intervals, $\sigma_{EU}^2 = 2.92$.

For the moderately skewed distributions, the probability values for each interval (1 = 0.01, 2 = 0.09, 3 = 0.15, 4 = 0.25, 5 = 0.35, 6 = 0.15) were adapted from Kozlowksi and Hults (1987). These proportions are also similar to those used in a recent investigation of interrater agreement measures (Burke et al., 1999). However, the weightings assigned to the intervals were redistributed slightly from the original values, since it appears that Kozlowksi and Hults (1987) overestimate the proportion of responses expected in the lowest interval of the six point scale.[7] For the skewed distributed, $\sigma_E^2 = E([X - E(X)]^2$ (James et al., 1984). E(X) can be calculated as: $\sum a_i p_i = 1(0.1) + \ldots + 6(0.2) = 4.29$, whereas $E([X - E(X)]^2 = (1-4.1)(0.1)^2 + \ldots (6-4.1)(0.2)^2 = 1.47$.

For the triangular distribution, $\sigma_E^2 = \frac{A^2 + 2A - 2}{24} = 1.92$ (see James et al., 1984).

Table 18 lists the average, maximum and minimum observed values for team r_{wg} for each survey scale, as well as the percentage of teams with values greater than 0.7 for the scale. In Table 18, r_{wg_u} denotes the r_{wg} values calculated using the uniform distribution as the expected response distribution, which, as noted, may represent an practical "upper bound" on r_{wg} if there is cause to suspect response bias in the data. The second set of values r_{wg_c}, denote the values calculated using the theoretical expected null distribution that represents the hypothesized direction of response bias, should it occur (see the discussion above). It should be noted here that in instances where the observed variance exceeded the expected variance r_{wg} was set to 0.0, as prescribed by James, Demaree and Wolf (1993, 1984).

As can be seen from Table 18, for all survey scales, the majority of team r_{wg_u} scores exceeded 0.7, which is the commonly accepted threshold for demonstrating within group agreement (George, 1990; Klein & Kozlowski, 2000; Van Mierlo et al., 2006) and the average r_{wg_u} scores were also greater than 0.7. In addition, for eight out of the 10 survey scales, more than 90% of team r_{wg_u} values also exceeded 0.7. The two remaining scales, *Goal Difficulty* and *Action Orientation*, had 88% and 75% of team r_{wg_u} values greater than 0.7, respectively.

The r_{wg_c} scores are, as expected, noticeably lower than the r_{wg_u} scores, but still, in general, support aggregation. This suggests that the conclusion to aggregate based on interrater agreement is fairly robust. All survey scales except for one had a majority of r_{wg_c} scores greater than 0.70 (the other scale, *Action Orientation*, had 45% of team

[7] The original proportions in Kozlowksi and Hults (1987) are as follows: 1= 0.1, 2 = 0.1, 3 = 0.1, 4 = 0.2, 5 = 0.3, 6 = 0.2. As Kozlowski and Hults (1987) note, an infinite number of theoretical skewed null distributions could be posited; however, except for the overweighting of the lowest two intervals, the distribution Kozlowski and Hults (1987) seems to match well what might be expected under a moderate social desirability bias.

values greater than 0.7). In addition, eight out of 10 survey scales had average r_{wg_c} scores greater than 0.7. However, two survey scales had average r_{wg_c} scores somewhat less 0.7. One of the two scales, *Goal Difficulty*, had a mean score of 0.65 and the other scale, *Action Orientation*, had a mean score of 0.6. These were the two survey scales where a triangular distribution was used to calculate r_{wg_c}, since there did not appear to be any reason to suspect social desirability effects – i.e., a skewed distribution. In general it appears there is less reason to suspect any response bias for these two scales than there is for suspecting bias – i.e., social desirability effects – in the other scales, so r_{wg_c} is likely a very conservative measure for these scales especially. It is also noted that *Action Orientation* is a slightly lower reliability measure than the other scales (see Table 16), with a Cronbach's alpha value of 0.64, and also contained only two items. Both of these characteristics would tend to lower observed r_{wg} values. James, Demaree and Wolf (1984) describe how scales having a large number of items tend to have higher r_{wg} values, since r_{wg} is based on the average within-team standard deviation across all survey items. As the number of scale items increases, the influence of any particular item on the average standard deviation – and therefore r_{wg} – decreases.[8] For scales with only a few items, fluctuations of team member responses by more than one response interval on one of the questions could have a fairly large impact on the average standard deviation, resulting in a noticeably lower r_{wg_c}, particularly for small teams. In fact, r_{wg} values are slightly attenuated in general for small groups – i.e., less than 10 members – when there is less than perfect agreement (Kozlowski & Hattrup, 1992).

Table 18. Interrater Agreement Values for Survey Scales

Scale	*Average* r_{wg_u}	Max r_{wg_u}	Min r_{wg_u}	% r_{wg_u} > 0.7	*Average* r_{wg_c}	Max r_{wg_c}	Min r_{wg_c}	% r_{wg_c} > 0.7
Goal Clarity	.882	.977	.072	94%	.702	.951	.000	69%
Goal Difficulty	.833	.969	.308	88%	.649	.950	.000	63%
Affective Commitment to Change	.935	1.000	.249	96%	.823	1.000	.000	92%
Internal Processes	.966	.997	.865	100%	.903	.993	.426	94%
Action Orientation	.756	.978	.000	75%	.595	.966	.000	45%
Management Support	.878	1.000	.000	90%	.705	1.000	.000	73%
Team Autonomy	.881	.980	.000	98%	.737	.969	.000	75%
Task KSA	.975	.998	.939	100%	.933	.996	.767	100%
Impact on Area	.908	.988	.497	96%	.751	.976	.000	69%
Attitude	.904	.990	.350	96%	.778	.980	.000	73%

[8] This is likely one reason that the two largest scales in the research – Internal Processes and Task KSA – had the largest average and minimum r_{wg} values.

Overall, the high r_{wg} values, combined with significant ICC(1) values, strengthen the argument for aggregation for the survey scales. Although the average r_{wg_c} values for *Goal Difficulty* and *Action Orientation* were somewhat less than 0.70, both of these variables had significant and "large" ICC(1) values – i.e., greater than 0.2 – and average r_{wg_u} values greater than 0.7, indicating strong support for aggregation in terms of ICC(1) and r_{wg_u}. In addition, as noted the r_{wg_c} values are intended to provide a practical "lower bound" on true agreement and are likely quite conservative for these two measures in particular (see discussion above).

In the case of the *Goal Clarity* variable, the r_{wg} values – both r_{wg_c} and r_{wg_c} – clearly support aggregation due to the similarity of ratings between individuals within the same team – i.e., both the average r_{wg_u} and the average r_{wg_c} are greater than 0.7. As previously described, the ICC(1) results for *Goal Clarity* are less supportive of aggregation. The ICC(1) value for *Goal Clarity* was significant only before a Bonferroni correction and ICC(1) was slightly less than 0.1, which is a recommended rule of thumb threshold justifying aggregation. Overall, these results indicate that *Goal Clarity* does not appear to vary as much between versus within teams as other variables. As mentioned previously, nothing theoretically requires a variable that is group-level – i.e., a true characteristics of the group, rather than individuals in the group – to vary between the groups in a given sample. However, such variables would not be useful in the analysis. Although the decision was made to aggregate *Goal Clarity* to the team-level based on the high r_{wg} scores and the relatively low p-value for ICC(1), one might posit that it is less likely to be a significant predictor in the regression analyses (see Chapter 4), due to the more limited variation between groups.

Following the conclusion that team-level aggregation was justified for the survey scales, team-level averages were calculated for each survey scale. At this time, in addition to examining the distribution of r_{wg} scores for each variable, each team was also examined to determine whether any teams displayed low r_{wg} scores – i.e., less than 0.7 – on all or most variables, which might suggest the team be excluded from further analyses due to lack of within-team consensus – i.e., reliable average measures. For this analysis, the more traditional r_{wg_u} measure was used, since r_{wg_c} provides a conservative "lower bound" on agreement and it is desirable to avoid losing information by spuriously removing teams, including information provided through additional variables not measured through the survey scales. None of the teams displayed low agreement on more than half of the variables. Most teams had zero or one variables with low r_{wg_u} scores – i.e., less than 0.7. Three teams had two variables with r_{wg_u} scores less than 0.7, one team had three variables with r_{wg_u} scores less than 0.7, and one team had five variables with r_{wg_u} scores less than 0.7. The researcher considered excluding the team with low agreement on five variables – i.e., 50% of the survey scales – from further analyses, since this was close to the cutoff criterion of more than half of the variables. However, based on the high agreement on the other five survey variables and the fact that this research includes several additional non-survey measures – thus the five affected survey variables were a minority of the total variables studied – the decision was made to retain this team for further analyses. It must be recognized, however, that for the five affected variables, the mean for this team is a less reliable representation of group-level opinion than it is for most of the other teams in the study.

3.8 Screening of Aggregated Variables

The final step in the data screening analysis, before testing the study hypotheses, was to analyze the distributions of the team-level variables, to determine whether the variables appeared to be normally distributed. Although parametric analysis methods are relatively robust to violations of the assumption of normality (Neter et al., 1996), variables that are strongly non-normal need to be transformed or analyzed using non-parametric methods. This analysis included not only the aggregated survey variables – *Goal Clarity*, *Goal Difficulty*, *Affective Commitment to Change*, *Internal Processes*, *Action Orientation*, *Management Support*, *Team Autonomy*, *Task KSA*, *Impact on Area*, *Attitude* – but also the variables collected only at the team-level via the Event Information Sheet – *Team Leader Experience, Work Area Routineness, Event Planning Process, Overall Perceived Success, Team Functional Heterogeneity, % of Goals Met, Tool Appropriateness* and *Tool Quality.*

Most of the survey variables appeared to be relatively normally distributed, and for all but three – *Goal Difficulty, Action Orientation* and *Internal Processes* – formal tests of normality were not rejected. Despite the rejection of the formal normality test, *Action Orientation* also appeared to be fairly symmetrically distributed, although negatively skewed (i.e., toward the lower end of the scale). *Internal Processes* appeared to be very symmetrically distributed, but had one low outlier. Similarly, *Goal Difficulty* was also fairly symmetrically distributed but skewed toward the low end of the survey scale with two low outliers. Thus, it appears that, in the case of these three variables, the departures from normality are not extreme and parametric analysis methods can be used. It should be noted, however, that the outliers on *Goal Difficulty* and *Internal Processes* may end up being influential cases in the analyses.

Of the non-survey measures, *Team Functional Heterogeneity* was the only variable where formal tests of normality were not rejected. Several other variables demonstrated only mild departures from normality. The distribution of *Overall Perceived Success* was relatively symmetric, but negatively skewed (i.e., toward the low end of the response scale) with three low outlier values. *Work Area Routineness* and *Tool Appropriateness* were similarly negatively skewed. *Tool Quality* was negatively skewed and also appeared more noticeably truncated than the other skewed distributions (i.e., a visual examination of the distribution suggests that, in an unconstrained response situation, the high tail of the distribution would have carried on beyond the maximum response scale value of 6.0). In general though, since departures from normality did not appear to be extreme, parametric analysis methods were used for the hypothesis tests.

However, four of the non-survey variables -- *% of Goals Met*, *Team Leader Experience*, *Team Kaizen Experience* and *Event Planning Process* – were more severely non-normal (i.e., severely skewed) and required transformation. A logarithmic (base ten) transformation was used and resulted in much more symmetric distribution, although in most cases, formal tests of normality were still rejected (*Team Kaizen Experience* was the exception). For *Event Planning Process* and *Team Kaizen Experience*, the transformation resulted in a variable that was fairly symmetrically distributed, although still somewhat positively skewed (i.e., toward high values), particularly for *Event Planning Process*. However, for *% of Goals Met* and *Team Leader Experience*, the distribution tightened but still remained highly skewed, since the mode for each of these variables was equal to the maximum or minimum value of the response continuum. In the case of *Team Leader Experience*, which measures the total number of Kaizen events the team leader has led or co-led, the mode was equal to the minimum response value (i.e., 1 or zero in the transformed version). Similarly, *% of Goals Met* had a mode equal to the maximum response value of 1.0 (i.e., 100%) in the untransformed scale and zero in the transformed scale. Since, for one team, the original *% of Goals Met* value was 0%, a small value (i.e., 1%) was used to replace this value in the log transformation, since the log of zero is undefined.

See Appendix S for a summary of the mean, median, maximum, minimum and standard deviation of the final set of study variables.

CHAPTER 4: RESULTS

4.1 Overview of Models Used to Test Study Hypotheses

The modeling process used to test the study hypotheses was as follows:

1. To test H1, H3 and H4, regression analyses were performed using generalized estimating equations (GEE) to account for (potentially) correlated residuals within organizations. The regression models were used to calculate the correlation coefficient for each variable pair of interest and testing the correlation for significance.
2. H2 was tested using ICC(1).
3. To test H5 – H8, regression analyses were performed using generalized estimating equations (GEE) to account for (potentially) correlated residuals within organizations. The regression models were used to determine which event input factors and event process factors had the most significant impact on each outcome in the overall population.
4. To test H9 – H10, for each outcome variable, mediation models were analyzed for any significant event process predictors. The purpose of the mediation analysis was to determine whether any event input predictors exhibited relationships with both the event process predictor and the outcome variable that are consistent with the mediation hypothesis – i.e., the hypothesis that the event input factor had an indirect effect on the outcome variable through the process predictor.

The relationships between event input factors, event process factors and outcomes were analyzed using regression modeling. As has already been mentioned, the current research includes two levels of nested data – i.e., lower-level units (teams) hierarchically nested within higher-level units (organizations). Thus, it is likely that lower-level observations within the same higher-level unit – i.e., teams within a given organization – may be correlated due to contextual factors.

In hierarchically nested data, within-group correlations – i.e., correlated residuals – can result in errors in the estimation of the standard errors of regression parameters in ordinary least squares (OLS) regression modeling, although the parameter estimates themselves remain asymptotically unbiased (Hox, 1994; Lawal, 2003). In particular, within-group correlation can result in substantial underestimation of standard errors – and thus an increase in Type I error for tests of regression parameters – if correlation within groups is strong and variation between groups is large. Thus, as in other modeling situations involving correlated error terms, if the correlation between groups is not taken into account, inference errors can result.

Hierarchical linear modeling (Raudenbush & Byrk, 2002), also referred to as “multilevel modeling,” “mixed effects models,” and “random coefficient modeling” (Davison et al., 2002) is one appropriate analysis technique for studies with nested data. Unlike OLS regression modeling, which assumes fixed intercepts and slopes across groups, HLM allows researchers to decompose the residual variance into variance components due to variance in intercepts across groups, variance in slopes across groups, and individual-level residual variance within-groups, (Bliese & Hanges, 2004), thus accounting for group effects within the residuals – i.e., correlated error terms. These variance components can then be tested for significance. Significant results indicate the need to include additional group-level variables – a classification variable for group can be used to model differences in intercepts, whereas additional group-level predictors will be necessary to account for differences in slopes. HLM allows researchers to develop a series of equations with predictors at each additional, higher-level to test cross-level effects.

HLM relies on the ability to model random, as well as fixed effects, and can be executed in SPSS using the “Mixed Models” procedure (see Peugh & Enders, 2005 for a tutorial on how HLM may be executed in SPSS), SAS using PROC MIXED (see Singer, 1998 for a tutorial on how HLM may be executed in SAS) or a number of specialty software packages (see Bliese & Hanges, 2004 for examples). However, HLM requires a fairly large sample size both across and within groups (James & Williams, 2000). Like most modeling methodologies, there are no absolute rules for the minimum sample size. Raudenbush and Byrk (2002) suggest roughly 10 observations per predictor per level – i.e., for a model with one predictor, 10 groups and 10 individuals within each group would be sufficient. However, for lower-level models – i.e., models with individual-level predictors only – the sample size requirement depends the suspected intraclass correlation – slightly fewer higher level observations can needed for testing lower level predictors if the intraclass correlation is expected to be small. Other sources have suggested 30 groups with 30 individuals each as a sufficient sample size for testing cross-level effects (Bassiri, 1998; Van der Leeden & Busing, 1994). The group (organizational) sample size in the present research was not large – i.e., six organizations – and the number of individuals (teams) studied within each organization varied from four to 15. Thus, it appeared that the current sample size both across and within organizations would be insufficient to support HLM. Indeed, test team-level (“level-1”) HLM models were executed in both SAS PROC MIXED and SPSS

“Mixed Models” and the maximum likelihood algorithm failed to converge, indicating that concerns about sample size were well-founded.

Another approach that is commonly used to correct for correlation between observations is generalized estimating equations (GEE) (Liang and Zeger, 1986). GEE is a robust estimation procedure specifically designed to account for correlated error terms – i.e. non-independence between rows – in regression. GEE is used with generalized linear models – i.e., regression models where the response variable can be specified as coming from any distribution within the exponential family (Horton & Lipsitz, 1999). GEE calculates parameter and standard error estimates based on the marginal distribution of the response variable by incorporating a correlation matrix to model the association between observations – e.g. measurement times, individuals, teams – within a given higher-level group or cluster – e.g., subjects, households, organizations. GEE uses an iterative generalized least squares approach to update parameter estimates based on observed correlations between residuals until convergence is achieved (Lawal, 2003). GEE generally results in very similar parameter estimates to OLS, since both estimation methods are asymptotically unbiased – i.e., they both approach the same value. The major difference is generally in the estimates of standard errors produced by the two methods. As mentioned above, the standard error estimates using OLS are biased when there is correlation between residuals for observations within the same group (Hanley et al., 2003). One difference from HLM is that GEE cannot accommodate differences in both slopes and intercepts across organizations – i.e., it is a single-level analysis method rather than a multi-level analysis method (Hanley et al., 2003). By incorporating the correlation between observations (teams) within a given cluster (organization), GEE effectively models differences in intercepts across cluster. If an exchangeable correlation structure is assumed, GEE is equivalent to a random effects model with a random intercept per cluster (Horton & Lipsitz, 1999). However, these random intercepts are not reported in GEE output, although the observed intraclass correlation is.

The GEE procedure requires the user to specify a form for the working correlation matrix used to develop estimates of parameter standard errors. GEE can be used to produce two sets of standard error estimates. One set of standard error estimates is “model based” – i.e., based on the working correlation matrix in the final iteration of the estimation. The second set of standard errors is “empirical” – i.e., based on a “robust” estimation procedure which employs a “sandwich” estimator. (Note, the actual parameter estimates are identical for the two standard error estimation procedures). If samples are relatively large – i.e. larger than twenty clusters – the “robust” or “empirical” standard error estimates produced by GEE are robust to misspecification of the working correlation matrix (Chang, 2000), and different specifications produce similar standard error estimates. This is due to the “robust” or “sandwich” variance estimator used to estimate the standard errors of regression parameters (Pan & Wall, 2002). (Drum and McCullagh (1993) point out that this variance estimator is actually more consistent than “robust” in the usual sense). However, in samples involving a relatively small number of clusters, the “empirical” estimates of parameter standard errors are less stable (Breslow, 1990; Drum & McCullagh, 1993; Cologne et al., 1993; Pan & Wall, 2002; Lawal, 2003; Hanley et al., 2003). In particular, the reported “empirical” standard errors may underestimate the true variance of the parameters – i.e., result in Type I error (Pan & Wall, 2002). For small samples at the cluster level, GEE standard error estimates can suffer from similar types of Type I error problems to those created through unmodeled residual correlation – i.e., use of OLS in the presence of correlated residuals. The inflated Type I error rate is somewhat problematic for the test of individual regression parameters using GEE “empirical” standard errors. However, it is much more problematic for the Wald test often applied to test the overall significance of the regression. For instance, a simulation study by Pan and Wall (2002) found that, for 10 groups with 20 individuals each, Type I error rates were roughly twice the nominal alpha for tests of individual regression parameters using the “empirical” standard error estimates (when applied to a logistic regression model with one predictor); however, they were roughly five times the nominal alpha for the Wald test applied to simultaneously test the significance of three predictors, again using the “empirical” standard error estimates. In general, for smaller overall samples, Type I error rates are more problematic for smaller cluster-level samples of larger clusters than for larger cluster-level samples of smaller clusters (Lawal, 2003).

Several approaches have been proposed to correct for inflated Type I error for the GEE “empirical” standard error estimates in smaller samples, either through direct modification of the sandwich estimator or modification of the significance test (e.g., Pan & Wall, 2002; Pan, 1999; Emrich & Piedmonte, 1992; Gunsolley et al., 1995). However, these approaches are computationally difficult – i.e., they are not implemented in mainstream statistical software and require direct manipulation of the sandwich estimator variance matrix, V_s – relatively new – i.e., do not appear to have been extensively used in applied research, and have been developed within the framework of logistic regression (although they appear equally applicable to other types of regression – see Breslow, 1990). Another commonly proposed approach for dealing with Type I error is examination of the “model based” standard errors as well as the “empirical,” or “robust,” standard errors. If the correlation structure – i.e., the working correlation matrix – has been correctly specified, the model based standard errors should provide more accurate estimates of the significance of the regression coefficients for small samples (Prentice, 1988). For instance, Lawal (2003) suggests that, if the number of clusters is less than 25, the analyst should focus on correctly specifying the working correlation matrix structure and report the "model based" standard errors, since the "empirical" standard errors do not provide a good estimate (Lawal (2003) cites unpublished course materials by Lipsitz (1999)). Similarly, for studies with fewer than 20 clusters, Horton & Lipsitz (1999) recommend using the "model based" standard errors instead of the "empirical" standard errors. Finally, Drum and McCullagh (1993) recommend examining and comparing results from both "model based" and "empirical" standard error estimates for small samples. It should be noted here that "correct" specification of the correlation matrix does not require the analyst to numerically specify the level of correlation between observations within a cluster – this is automatically and iteratively calculated through the GEE procedure. What is required is that the analyst specify a reasonable overall structure for the pattern of correlation within clusters – i.e., exchangeable/compound symmetry, unstructured, autoregressive, independence, etc.

In the current model, compound symmetry – i.e., an exchangeable matrix – appears to be the most reasonable assumption for the working correlation structure. Compound symmetry assumes that the correlation between all observations (teams) within a given cluster (organization) are equal, and the same correlation structure is applied to all clusters (Hardin & Hilbe, 2003). This structure is recommended for studies where there appears to be no natural orderings of observations (Horton & Lipsitiz, 1999; Hardin & Hilbe, 2003). In the context of the current study, the use of an exchangeable correlation matrix is equivalent to assuming that there will be some intraclass correlation between observations (teams) within a given cluster (organization) due to contextual effects.

Other options for modeling the correlation structure included autoregressive (and similar time-based structures), unstructured and independence. Autoregressive structures assume that observations that are closer together in time have stronger correlations than observations that are further apart in time. In the current research, some argument could be made for treating the correlation structure as time ordered, because each event is associated with specific calendar dates. However, upon closer examination, time effects are not expected to be a particularly strong source of correlation between observations. First, since the boundary criteria for selecting organizations specified that the organizations had been conducting Kaizen events for at least one year prior to enrolling in the study, large correlations between outcomes for adjacent observations due to an organizational learning curve are not expected. Second, the time window between observations is generally so small that observations can be considered functionally concurrent in terms of organizational climate and culture variables – i.e. contextual variables – which only vary on a much greater time scale. In fact, in one organization, some of the events were actually concurrent, occurring during the same week. In contrast, the effects of organizational climate and organizational culture variables are expected to be relatively stable over the relatively short study window – i.e., six to nine months. Third, the observations are not equally spaced across organizations, thus it seems unlikely that the same pattern of correlation between adjacent observations would hold across organizations. Finally, it appears that greater than average similarity between any two observations within a given cluster would be more likely to be due to similarity of focus – i.e., event design – rather than time effects. Thus it does not appear that autoregressive or other time based correlation structures would be appropriate.

"Independence" specifies a working correlation matrix in which all pairwise correlations are assumed to be zero. Although this structure produces appropriate "empirical" standard error estimates when samples are large, it seems inappropriate for calculating "model based" standard errors in this context, since it assumes no correlation between observations. Finally, in an unstructured matrix the correlation between each pair of observations would be estimated separately; however, the same correlation structure would be applied to all clusters (organizations). This seems more likely to result in error that the compound symmetry assumption, since there is no reason to believe that the same pattern of unique pairwise correlations holds across clusters (organizations), since the observations are not equally spaced in time or otherwise equally ordered. In addition, an unstructured correlation matrix may be problematic for unbalanced data, such as the current research (Hardin & Hilbe, 2003). Thus, as suggested in the GEE literature, an exchangeable correlation matrix – i.e., assuming that all pairs of observations within a given cluster are equally correlated – appears the most reasonable assumption for the current research, given the lack of natural ordering of observation and the expected presence of contextual effects (i.e., intraclass correlation).

4.2 Analysis of H1 - H4

H1, H3 and H4 were tested by: 1) calculating the Pearson correlation coefficient for each variable pair of interest and 2) testing the associated regression relationship for significance. GEE were used to account for the nested structure of the data – i.e., teams within organizations. The correlation coefficient was calculated as the square root of the coefficient of determination for the regression model where one of the two outcomes was regressed on the other. Due to the nested structure of the data, the correlation coefficients was not tested for significance directly – i.e., using a t-test – since the degrees of freedom for the test could not be assumed to be correct. Instead, the regression coefficient $\widehat{\beta}_{GEE}$ was tested for significance, which provides conceptually equivalent results (Montgomery & Runger, 1999).[9] Due to the fact that the correlation coefficients were calculated using GEE instead of OLS, the correlation coefficients were sometimes slightly different for the regression of Y on X versus X on Y – whereas in OLS the correlation coefficients are identical for both regressions. Thus, the minimum correlation coefficient was used as the estimate of the correlation between X and Y. The maximum p-value from the two regressions (Y on X and X on Y) was used for the significance test, because, under GEE, the significance of the two regression coefficients can also differ. However, it should be noted that in no cases were the regression coefficient p-values so different that they would have implied different results. Also, because GEE was used, regression coefficient p-values do not directly coincide with the magnitudes of the correlation coefficient when comparing across pairs. When comparing across pairs of variables, a smaller correlation coefficient for one pair compared to another does not necessarily indicate a smaller p-value for the associated regression.

There are ten pairwise comparisons of interest:

- Social system outcome analysis (two variables, one comparison) -- *Attitude* (AT) and *Task KSA* (TKSA)
- Technical system outcome analysis (three variables, three comparisons) -- *% of Goals Met* and *Overall Perceived Success* (OVER), *% of Goals Met* and *Impact on Area* (IMA), and OVER and IMA
- Social system and technical system outcome comparison (five variables, six comparisons) – AT with IMA, AT with OVER, and AT with *% of Goals Met*; and TKSA with IMA, TKSA with OVER, and TKSA with *% of Goals Met*.

Because the total number of pairwise comparisons to be made is large (10), a Bonferroni correction is used to adjust the alpha value for the number of planned comparisons. Since the desired family confidence level is 0.05, an alpha value of 0.05/10 = 0.005 is used in the hypothesis tests. Table 19 shows the pairwise correlations for these 10 relationships. The maximum p-values for the two regressions involving the pairwise relationship (Y on X and X on Y) are shown in the second half of the table, below the correlations. Correlations with an asterisk are associated with regressions that are significant at a 95% family confidence level. Appendix T displays the full results of the analysis – i.e., the two regressions and two p-values for each pairwise relationship.

[9] In fact, in OLS, testing the regression coefficient for significance provides directly equivalent results to testing the correlation for significance.

While H1, H3 and H4 were tested through correlation analysis, H2 ("Social system outcomes will occur primarily at the team level, rather than individual level, indicated by significant intraclass correlation for social system outcome variables") was tested through an examination of ICC(1). Although demonstrating a significant ICC(1) was important for all survey variables in terms of providing justification for aggregation to the team level, demonstrating a significant ICC(1) is of additional theoretical importance for the social system outcomes – i.e., AT and TKSA. As described in Chapter 3, learning theory suggests that group learning outcomes occur at the group level rather than individual level – i.e., are shared group properties. Thus, demonstrating a significant ICC(1) was of particular theoretical importance for the social system outcomes in order to demonstrate alignment with learning theory. As indicated in Table 17, the ICC(1) values for AT and TKSA were both significant, even when a Bonferroni correction is employed ($p < 0.0001$ for both values). Thus, H2 is supported. Table 20 shows the study hypotheses, the null hypotheses and the test results. Results are discussed in detail in Chapter 5.

Table 19. Pairwise Correlations for Outcome Variables and Regression Significance Tests

Correlation	AT	TKSA	OVER	IMA	% of Goals Met
AT	1				
TKSA	.710*	1			
OVER	.122	.155	1		
IMA	.632*	.688*	.224	1	
% of Goals Met	.000	.000	.163	-.019	1
Max. Sig. of $\widehat{\beta}_{GEE}$					
AT	-				
TKSA	.0000	-			
OVER	.4068	.3326	-		
IMA	.0000	.0000	.1073	-	
% of Goals Met	.9937	.3451	.2982	.5882	-

Table 20. Study Hypotheses and Test Results

Hypothesis	Comparisons (α = 0.005)	Result
H1: Social system outcome variables will be significantly correlated at the team level. H_0: Social system outcome variables are not significantly correlated at the team level.	• AT and TKSA, $r = 0.710$, $p < 0.0001$	Supported
H2: Social system outcomes will occur primarily at the team level, rather than individual level, indicated by significant intraclass correlation for social system outcome variables H_0: The intraclass correlation for social system outcomes is not significant	• AT, ICC(1) = 0.300, $p < 0.0001$ • TKSA, ICC(1) = 0.121, $p < 0.0001$	Supported
H3: Technical system outcome variables will be significantly correlated at the team level. H_0: Technical system outcome variables are not significantly correlated at the team level.	• IMA and OVER, n.s. • IMA and % of Goals Met, n.s. • OVER and % of Goals Met, n.s.	Not Supported
H4: Social system outcomes will be significantly correlated with technical system outcomes at the team level. H_0: Social system outcomes are not significantly correlated with technical system outcomes at the team level.	• AT and IMA, $r = 0.632$, $p < 0.0001$ • TKSA and IMA, $r = 0.689$, $p < 0.0001$ • AT and OVER, n.s. • TKSA and OVER, n.s. • AT and % of Goals Met, n.s. • TKSA and % of Goals Met, n.s.	Partially Supported

4.3 Regression Analysis to Test H5 – H8

4.3.1 Screening Analysis Prior to Building Regression Models

The regression to be performed in this research is exploratory in the sense that it is not known which of the 14 measured independent variables are most strongly related to each of the five measured outcomes. Thus, for each outcome variable, a selection procedure was used to iteratively narrow the set of independent variables to a set of the most significant predictors. Prior to creating the regression models, the input data – i.e. response/dependent variables and predictor/independent variables – were examined through a screening process to determine whether there appear to be any extreme violations of basic assumptions of linear regression, which would impact the results of the model building procedure. For instance, as described in Chapter 3, an exploratory analysis of the team-level variables revealed that most of the outcome variables appeared relatively continuously distributed, although this was less true of *% of Goals Met* and *Overall Perceived Success* than the other outcome variables. In the case of non-continuous outcome variables, regression can still be performed but another response distribution must be used – e.g., logistic regression. However, given the response variables are at least somewhat continuous, it is preferred to first attempt to model the *% of Goals Met* variable as continuous, rather than to lose information by transforming the variable into a binary or categorical form. Follow-up post-hoc analyses could then be used to examine a binary or categorical version of this variable to determine whether the inferences differ from the model performed using normal (Gaussian) regression.

In addition to examining the outcome variables, the degree of correlation between the independent (predictor) variables was examined to determine whether multicollinearity was likely to be a problem in the present research. It is commonly recognized that multicollinearity – i.e., a high degree of correlation between one or more predictor variables – can create a number of problems in regression, including instability of the regression solution (i.e., non-uniqueness of the solution and parameter estimates) and inflated standard error estimates for regression parameters (Neter et al., 1996).

In OLS, the variance inflation factor (VIF) is interpreted as the extent to which the variance of the regression parameter for the *k*th predictor is inflated when the other p – 2 variables[10] are in included in the regression (Neter et al., 1996). In GEE, it is not clear that this interpretation strictly applies, due to different methods of estimating the standard errors of the parameters, and the GEE estimation procedure in SAS PROC GENMOD does not include an option for generating VIF. However, the VIF still appears to be an appropriate measure of the extent to which a given predictor covaries with all the other predictors in the model – i.e., the extent to which the variance of the given predictor can be predicted by the other predictors – and thus a useful tool for diagnosing multicollinearity. This is based on the fact that the actual parameter estimates attained through OLS and GEE regression are very similar, because. both are asymptotically unbiased. Thus, the residuals produced by both models are also very similar. The VIF is calculated using Equation 4 (Neter et al., 1996), where R_k is the R value that results when a given predictor is regressed on the other n - p predictor variables:

$$VIF_k = \frac{1}{1 - R_k} \qquad \textbf{(4)}$$

As Equation 5 shows, *R* is derived from the residuals – $\sum(Y_i - \hat{Y}_i)$ – as well as the deviation around the grand mean -- $\sum(Y_i - \overline{Y})^2$ -- both of which would be similar under OLS and GEE estimation.

$$R = \sqrt{R^2} = \sqrt{\frac{SSR}{SSTO}} = \sqrt{1 - \frac{SSE}{SSTO}} = \sqrt{1 - \frac{\sum(Y_i - \hat{Y}_i)^2}{\sum(Y_i - \overline{Y})^2}} \qquad \textbf{(5)}$$

---

[10] Note, p represents the total number of parameters being estimated in the model and one of the parameters being estimated is the intercept.

Likely due to similar reasoning, other studies using GEE have also examined VIF as an aid for diagnosing multicollinearity (e.g., Pickering & Kisangani, 2005; Whitford & Yates, 2003; and Szinovacz & Davey, 2001). Table 21 lists the VIF for each predictor generated using PROC REG in SAS 9.1.3 (i.e. OLS), as well as the VIF calculated from PROC GENMOD (i.e., GEE) output. A commonly accepted rule of thumb for interpreting VIF is that an individual VIF value of 10 or greater indicates a problem with multicollinearity, as does an average VIF substantially greater than three. In the present example, the maximum observed VIF was less than 3.0, and the average VIF was less than 2.0. Thus, it appears that multicollinearity is not severe in the current regression modeling.

Table 21. VIF for Predictor Variables

Predictor	**VIF_{OLS}**	**VIF_{GEE}**
Goal Clarity	2.96	2.91
Goal Difficulty	1.59	1.27
Team Autonomy	2.15	1.77
Team Functional Heterogeneity	1.39	1.22
Total Kaizen Experience	2.05	1.56
Team Leader Experience	1.63	1.61
Management Support	2.20	2.19
Event Planning Process	1.78	1.78
Work Area Routineness	1.75	1.70
Affective Commitment to Change	2.48	2.43
Action Orientation	2.29	2.14
Internal Processes	2.00	1.97
Tool Appropriateness	1.61	1.56
Tool Quality	1.58	1.54
Average VIF	1.96	1.83
Max VIF	2.96	2.91

4.3.2 Model Building Process

Following investigation of the VIF, and the conclusion that multicollinearity was not a severe problem in the current modeling process, regression models were built separately for each of the five outcome variables: the two social system outcomes, *Attitude* and *Task KSA*, and the three technical system outcomes, *Overall Perceived Success*, *% of Goals Met*, and *Impact on Area*. The model building process included three basic steps: 1) a manual backward selection process using GEE models; 2) an automated search process (backward selection and stepwise selection) using OLS models, because these automated procedures were not available for the GEE models; and 3) an all possible regression procedure to identify "good" subsets of predictors – the RSQUARE, ADJRSQ, CP and MAXR selection procedures PROC REG in SAS 9.1.3 – again using OLS because, again, these automated procedures were not available for the GEE models. The aim of this process was identifying convergence of the set of the predictor variables to be included in the final model.

The primary model selection procedure needed to be performed using GEE estimates, rather than OLS estimates, since tests of the significance of regression parameters are based on the standard errors of the parameters and the results of these OLS procedures could not be assumed to be accurate (Hanley et al., 2003). In addition, most automated selection procedures use a fixed alpha, instead of adjusting for family confidence depending on the number of predictors – i.e., a Bonferroni correction. However, if the standard errors of the parameters are not strongly biased, the results of OLS automated selection procedures and the GEE manual procedure would be expected to be similar. Thus OLS selection procedures can be examined as informal support for the GEE selection results and, as mentioned, to identify additional potential "good" subsets of predictor variables.

In the OLS automated search procedures in SAS PROC REG, the RSQUARE selection procedure reports variable subsets of size p – in the current study, p goes from 1 to 14 – in order of descending R^2. The ADJRSQ selection procedure reports variable subsets of any size in order of descending adjusted R^2 – i.e., R^2 adjusted for the number of variables currently in the model (Neter et al., 1996). Similar to the ADJRSQ procedure, the CP selection procedure reports variable subsets in order of ascending C_P. The C_P criterion is a measure of the total mean square error of the fitted values from the given regression model, including error due to (assumed) random deviation and error due to regression model bias (Neter et al., 1996). The C_P criterion assumes that in the full model – i.e., the model containing all p – 1 candidate predictors – MSE provides an unbiased estimate of σ^2, the true residual variance; this implies that all the important variables related to the response have been identified and included in the full model. A "good" subset under the C_P criterion would have a C_P that is "small," indicating relatively lower total error, and that is roughly equal to p – i.e. the number of parameters in the model – indicating low bias in the regression model. Although these automated search procedures can be effective in determining which subsets of variables have the most predictive power – i.e., highest R^2 given the model size – they do not guarantee that the variables included in the "best subsets" are significant. This is particularly true since the significance levels will ultimately be based on GEE model results, which account for correlation between observations, instead of OLS results. Thus, the automated search procedures, particularly the "best subsets" procedures are used as a tool for testing the robustness of the backward selection model specified in stage one using GEE, rather than as a primary selection tool. Promising subsets were implemented in GEE for comparison versus the model identified through the manual backward selection process.

If previous research had established the theoretical order of importance of the predictor variables, a hierarchical regression procedure would generally be preferred to an exploratory selection process. In a hierarchical regression procedure, the most important known predictors from previous research are entered first, followed by less important known predictors, until a set of significant known predictors has been identified. Then, new predictors are entered to test their significance while controlling for known effects; again, if appropriate, these new predictors can be entered in order of their theoretical importance. The final model is then refined until all predictor coefficients are adequately significant, depending on the significance level selected by the researcher (Neter et al., 1996; Field, 2005). The preference for hierarchical entry versus search procedures is due to the fact that exploratory regression results are less likely to be replicable across studies. However, the current research is exploratory in the sense that it appears to be the first, empirical field study of the initial outcomes of multiple Kaizen events in multiple organizations. Therefore, due to lack of previous work in this area, there is no clear order of importance of the 14 candidate variables. All of the 14 candidate predictor variables were identified as important theoretical predictors of Kaizen event outcomes, based on previous studies of related organizational mechanisms (i.e., projects and teams – see Chapter 2) or the Kaizen event practitioner literature. Although, even within this set of 14 "important" candidates, the researcher could make some subjective judgments about the likely order of relative importance – i.e., due to frequency of mention in the practitioner literature or the outcomes of studies of related mechanisms – there is no clear groundwork for making these *a priori* assumptions about relative importance. Thus, adopting a hierarchical entry method in this research based on perceived order of variable importance could induce bias, particularly since the outcome variables measured in related organizational studies were somewhat different than the outcomes studied in this research. Thus, an exploratory variable selection approach was applied as the primary strategy in this study.

However, two "hierarchical" approaches will be considered as complements to the backward selection procedure containing all 14 candidate predictors. Arguably, the only variable completely new to this research is *Action Orientation*, which is cited as a unique feature of Kaizen events (see Chapter 2). The other variables have all been tested, in some form, in research on teams or projects in different contexts. What has not been tested, however, is the unique combination of the study variables in the context of Kaizen events; particularly their effect on the different outcomes studied in this research. Another hierarchical regression could be based on the division of predictor variables into event input factors and event process factors. Arguably, event process factors could be considered more salient than event input factors, since, as the research model (Figure 1) shows, event input factors are hypothesized to act indirectly on the outcomes through the event process factors – i.e., this is a mediation hypothesis – as well as directly in some cases. Therefore, if a hierarchical regression approach were to be used, two approaches appear most reasonable: 1) performing a backward selection procedure with all study variables entered except for *Action Orientation*, then testing the effect of adding *Action Orientation* to the models; and 2) testing process predictors first using a backward selection procedure and then testing event process predictors. In the second hierarchical regression procedure, none of the significant event process predictors will be removed until the final stages of model refinement – i.e., once all significant event input factors have been identified through the backward selection process – even if the observed p-value exceeds the observed p-values for all event input factors. As mentioned, these two approaches will be used as a complement to the backward selection procedure containing all 14 variables to evaluate the robustness of the initial solution. Again, the GEE "model based" standard errors will be used to implement these regression procedures. Future research can use the results of this modeling in hierarchical regression analyses to determine whether the findings hold across additional studies, particularly studies that sample events in additional organizations of markedly different characteristics, and to further refine the regression models – e.g., to add additional predictors, etc.

The GEE models were executed using PROC GENMOD in SAS 9.1.3. All dependent variables were modeled as normal – thus an identity link function was used – since the exploratory analysis of the team-level variables had revealed relatively continuous distributions of the outcome variables (see Chapter 3). However, as mentioned, *% of Goals Met* had a more truncated and less symmetric distribution than the other outcome variables and *Overall Perceived Success* had a symmetric but truncated distribution.

As mentioned, in smaller samples, the "robust" or "empirical" estimates of parameter standard errors are less stable and likely to be downwardly biased (Breslow, 1990; Drum & McCullagh, 1993; Cologne et al., 1993; Pan & Wall, 2002; Lawal, 2003; Hanley et al., 2003). However, if a reasonable correlation structure has been specified for the working correlation matrix, the "model based" standard errors provide better estimates of the true standard errors – and thus, significance – of the regression parameters. Thus, because the current research involved a small sample at the cluster (organization) level, both "model based" and "empirical" standard errors were examined; however, the "model based" standard errors are expected to be more trustworthy, provided the specification of the working correlation matrix is reasonable.

As mentioned previously, the GEE models were built through a backward selection procedure. A backward selection procedure was used instead of a forward or stepwise procedure, since a backward selection procedure is less likely to result in the exclusion of important variables than a forward or stepwise selection procedure (Neter et al., 1996; Field, 2005). All of the 14 candidate predictor variables were initially entered into the model. At each stage, the predictor with the lowest *p*-value was removed if and only if this *p*-value was less than 0.10/p, where p was the number of parameters being estimated in the current model – this is the Bonferroni correction for Type I error. A liberal, overall family confidence value of 0.10 was chosen since this would reflect an alpha of 0.05 for individual parameters in a model with two parameters – i.e., an intercept and a single predictor variable. Since the primary purpose of this regression is descriptive – i.e., identifying which independent variables are significant predictors of each dependent variable, rather than predictive – i.e., building the model that minimizes residual errors, it was deemed especially important to preserve strong evidence that the individual predictors included in the final model were significant. However, after the initial set of predictors was specified through the backward selection procedure, additional variables where $0.10/p < p < 0.05$ were considered for inclusion in the final model, although evidence for their significance as predictors is not as strong as those with smaller p-values. This is due to the fact that the purpose of the research is to identify which set of predictor variables appear to be related to each outcome variable, with the additional aim of identifying variables than can be further tested in future research. Thus, for the final models, preference was given to identifying larger sets of variables with slightly higher p-values in order to increase the explanatory power for researchers and organizations – i.e., there was a slight preference for Type I versus Type II error in the research. However, no variables with a p-value greater than 0.05 were considered for inclusion in the final models and final models were only adopted if at least roughly half the predictors also had a p-value less than or equal to 0.10/p.

As mentioned, both the "model based" standard error estimates and the "empirical" standard error estimates were examined in the backward selection procedure, although more credence was given to the "model based" standard error estimates since these were believed to be more accurate (see the discussion earlier in this chapter). However, where the two estimates diverged in terms of order of variable significance, the backward selection procedure was applied separately to each set of estimates to determine what final set of predictor variables would be implied by the “empirical” standard error estimates. The “model based” standard error estimates for the final set of predictor variables implied through the “empirical” standard error estimates were calculated and the overall fit of the regression model versus the model implied through the “model based” standard error estimates was examined. The next five sections now describe the results of the model building process for the each of the five outcome variables, respectively.

4.3.3 Model of Attitude

The backward selection procedure using GEE “model based” standard error estimates, the two hierarchical selection procedures and the automated backward and forward selection procedure using OLS converged upon the same set of three predictors: *Team Functional Heterogeneity*, *Management Support* and *Internal Processes*. In addition, this subset of predictors was the best three variable solution using the RSQUARE, CP and MAXR selection procedure results from SAS PROC REG. (Note, the ADJRSQ procedure only prints solutions with the highest ADJRSQ and no three variable solutions were displayed; although, as mentioned, it is likely that most of the printed solutions contain one or more variables that are nonsignificant).

The GEE “empirical” standard errors resulted in a similar, although slightly different, solution. Although *Management Support* and *Internal Processes* were included in the final solution – reinforcing their significance as predictors – *Team Functional Heterogeneity* was not. However, *Action Orientation* was included as a final predictor, resulting in a three variable solution. Overall, there was a lack of strong evidence for considering this alternative solution. Most importantly, the GEE “model based” standard error estimates for this model suggested that none of the predictors would be considered significant in the final solution. The p-values for all predictors were greater than 0.05 and therefore much greater than 0.10/4 = 0.025. Next, this subset of predictors was only the third best three variable solution listed in the RSQUARE selection procedure results using PROC REG, while the original solution was the best three variable solution. Finally, the adjusted R^2 for the OLS model was smaller for this model than for the model containing *Team Functional Heterogeneity*, *Management Support* and *Internal Processes* (0.555 versus 0.577).

The final solution, including the GEE parameter estimates, “model based” and “empirical” standard error estimates and R^2, and the OLS parameter estimates, standard error estimates and R^2, is presented in Table 22. The reported p-values are shown in parentheses after the standard error estimates. The label SE_{MB} refers to the “model based” standard error estimates using GEE, while the label SE_E refers to the “empirical” standard error estimates.

Other goodness-of-fit indicators that are often included in the reporting of GEE regression results include deviance statistics, log likelihood statistics, and Wald statistics (for the contrast to test the null hypothesis that all regression parameters are equal to zero). This Wald statistic is therefore analogous to the F-test for the overall significance of the regression used in OLS. Although an automatic part of PROC GENMOD output, deviance statistics and log likelihood statistics are not valid for GEE models, since GEE is a quasi-likelihood approach (SAS Institute Inc., 2006; Hardin & Hilbe, 2003). A "naïve likelihood ratio test" may be performed by using likelihood statistics from the associated independence models – i.e., those implemented through GLM without GEE – and adjusting the degrees of freedom for the test based on the "empirical" standard error estimates from GEE (Hardin & Hilbe, 2003). However, this approach requires the assumption that the "empirical" standard errors are unbiased, which is likely not true of this research, given the small sample size at the cluster (organization) level. In the test of regression coefficients, the Wald statistic has been shown to be overly liberal in general – i.e., prone to Type I error – particularly for smaller samples, since it is based on the "empirical" standard error estimates (Pan & Wall, 2002). Hardin & Hilbe (2003) present a corrected version of the Wald test that uses the "model based," rather than "empirical," standard error estimates. However, this test requires direct manipulation of the variance matrix – i.e., advanced matrix algebra – and is not available in standard software packages. Thus, the deviance, likelihood and Wald statistics were not used to assess goodness-of-fit in the current analysis. Instead, the analysis relies on the significance levels of individual predictor variables and the magnitude of R^2 and adjusted R^2 for both the OLS and GEE models. Although R^2 is not a standard part of SAS PROC GENMOD GEE output, Hardin & Hilbe (2003) particularly recommend calculating R^2 for GEE models.

Table 22. Final Regression Model for Attitude

	$\widehat{\beta}_{GEE}$	SE_{MB} $\widehat{\beta}_{GEE}$	SE_E $\widehat{\beta}_{GEE}$	$\widehat{\beta}_{OLS}$	SE $\widehat{\beta}_{OLS}$
Intercept	.467	.614 (.447)	.438 (.287)	.477	.627 (.450)
Team Functional Heterogeneity	-.547	.218 (.012)	.258 (.034)	-.551	.223 (.017)
Management Support	.250	.101 (.013)	.070 (.000)	.248	.102 (.019)
Internal Processes	.694	.115 (.000)	.063 (.000)	.693	.115 (.000)
OLS $R^2 = .602$; *OLS* $R_a^2 = .577$; $F_{3,47} = 24.01$*** *GEE* $R^2 = .605$; *GEE* $R_a^2 = .580$; $\rho = -0.019$					

As expected, the regression coefficients are very similar for the GEE estimation versus the OLS estimation, since both are asymptotically unbiased. In addition, the standard error estimates were quite similar, which is expected given the fact that the observed intraclass correlation reported by the GEE procedure was only -0.019, indicating that the correlation within clusters (organizations), given the final set of predictors, was not large. In fact, the fact that the intraclass correlation is actually negative in the final model suggests that more variation occurs between clusters (organizations) versus within clusters. However, it should also be noted that the intraclass correlation may not be significantly different from zero. Thus, the GEE model would be expected to give similar results to a model built under an independence assumption (i.e., OLS). The GEE "empirical" standard error estimates were, in general, smaller than the "model based" standard error estimates. These results were expected, since, as previously discussed, the GEE literature indicated that "empirical" standard error estimates might be downwardly biased due to the small sample size at the organizational level. The only exception was for the standard error estimates for the *Team Functional Heterogeneity* coefficient, which was slightly larger for the "empirical" estimate versus the "model based" estimate. It is important to note that all the regression coefficients in Table 22 would be considered statistically significant at the 0.10/4 = 0.025 level for all models, except for *Team Functional Heterogeneity* under the "empirical" standard estimation. Using the "empirical" standard error estimates, $0.025 < p < 0.05$ for the regression coefficient of *Team Functional Heterogeneity*. The implications of this solution are further discussed in Chapter 5.

As an analysis of model robustness, additional "good" subsets of predictors from the RSQUARE, ADJRSQ, CP and/or MAXR results were examined to help evaluate the robustness of the current solution. The underlying logic applied was as follows. Because the objective of the research was to identify variables with a significant relationship to outcomes, smaller – i.e., two or one variable – solutions were generally not considered as a replacement for the model identified through the backward selection procedure. However, some of the most promising larger – i.e., four or five variable – solutions were examined to determine whether all predictor variables were significant using GEE "model based" standard error estimates for any of these solutions. Particular attention was paid to solutions which contained the original solution found, but included one or two additional variables.

No larger solution where all predictors had p-value less than 0.05 were found.[11] However, one four factor solution nearly consisted of predictors that all had p-value less than 0.05. In the regression consisting of *Team Functional Heterogeneity*, *Management Support*, *Action Orientation* and *Internal Processes*, the p-values for all variables except for *Action Orientation* were less than or equal to 0.025. The p-value for *Action Orientation* was slightly greater than 0.07. This was the second best four variable solution using the RSQUARE and CP selection procedures. In addition, in both hierarchical selection procedures, *Action Orientation* represented the last variable

[11] Given the 14 predictor variables, an exhaustive search of all possible solutions was infeasible. Thus, there is always some chance that a "better" solution has been missed.

removed in the final stages of model refinement. In the first hierarchical procedure, *Action Orientation* could not be added to the specified model – *Team Functional Heterogeneity*, *Management Support*, and *Internal Processes* – because the p-value exceeded the established threshold of 0.10/p (p = 0.07). In the second hierarchical procedure, *Action Orientation* was removed from the model in the final stage of model refinement (again, due to p = 0.07), leaving Function Heterogeneity, *Management Support* and *Internal Processes* as the final set of predictors in the model. Finally, *Team Functional Heterogeneity*, *Management Support*, *Action Orientation* and *Internal Processes* was also the solution in the next to last step in the manual backward selection procedure using the GEE "model based" standard error estimates – i.e., *Action Orientation* was removed to yield the final solution – and represents a merging of the solutions from the "model based" and "empirical" standard error estimate selection procedures. Thus, it appears that there is some, albeit relatively weak, evidence that *Action Orientation* also has a small, positive effect on *Attitude* ($\widehat{\beta}_{GEE}$ = 0.088). However, due to the fact that the p-value for *Action Orientation* was greater than 0.05, this alternate solution was not ultimately adopted.

Following the selection of the final subset of model variables – *Team Functional Heterogeneity*, *Management Support* and *Internal Processes* – additional aspects of the model fit were examined to determine whether there appeared to be any serious errors of model specification. First, the residuals from both the GEE and OLS models were plotted against predicted values, and partial regression plots were also created. Similarly to OLS regression, residual analysis is recommended for regression models using GEE to detect effects such as nonlinearity, omitted variables, etc. (Chang, 2000; Hardin & Hilbe, 2003). None of the plots suggested any significant problems, such as departures from linearity, etc. In addition to the residual plots, an initial analysis of the residual distribution was conducted to determine whether the observed pattern of residuals appeared to be random. Chang (2000) suggests that, when GEE is used, in addition to the examination of residual plots, a non-parametric Wald-Wolfowitz run test should also be employed to determine whether there any patterns of systematic departure – i.e., bias – in the residuals. In an unbiased model, after GEE has been employed to account for the correlation structure, a random pattern of negative and positive residuals should be observed. There should not be many (long) series of consecutive positive or negative residuals. The Wald-Wolfowitz test calculates to what extent the number of series of positive versus negative residuals appears to depart from a random distribution – i.e., a significant result suggests a non-random pattern. The Wald-Wolfowitz test is specifically recommended for longitudinal, within-subject data to help determine whether the specification of the correlation structure is reasonable but also appears to be applicable to non time-dependent clustered data – e.g., teams within organizations. To conduct the run test, observations are ordered by organization – i.e., the same order that they were originally entered into SAS to create the GEE models. The Wald-Wolfowitz test was used on both the GEE-based and OLS-based residuals. The distribution of the OLS residuals was also examined using the cumulative normal probability plot generating using the SPSS "Linear Regression" procedure. Neither of the sets of residuals appears to depart significantly from normality (for both, $p = 0.20$).

Residuals were also analyzed to determine whether there were any influential cases – i.e., outliers. All standardized residuals were less than 3.0, indicating that there was no strong evidence of influential cases. In fact, all but three of the standardized residuals were less than 2.0. The standardized residual with the largest absolute value was –2.58. Other measures of influence were examined. For the OLS models, Cook's distance was examined. Cook's distance is a measure of the amount of influence a particular case exerts on the regression parameters (Neter et al., 1996) and can be evaluated through the F statistic with p and n-p numerator and denominator degrees of freedom, respectively. None of the Cook's distance values were significant ($p < 0.05$) or nearly significant ($p \leq 0.10$) for the final regression model, and the maximum Cook's distance was only 0.191 ($p = 0.942$).

Finally, the three-way and two-way interactions of the variables in the final mode were tested to determine whether there was evidence of significant interactions of these terms. None of the interaction terms were significant.

4.3.4 Model of Task KSA

In general, fitting the regression model of *Task KSA* was more complex than for the other two survey outcome measures – i.e., *Attitude* and *Impact on Area*. This could be due to the fact that *Task KSA* is a "large" variable measured by 10 different items and therefore many different event input and event process variables may be somewhat related to *Task KSA*, particularly when levels of other variables are held constant – i.e., suppressor effects (Field, 2005). However, there were some similarities between the solutions produced through each selection procedure. The manual backward selection procedure using GEE "model based" standard error estimates and the two hierarchical search procedures resulted in the following solution: *Team Autonomy* and *Internal Processes*. This was the second best two variable solution using RSQUARE. The manual backward selection procedure GEE "empirical" standard errors produced the following solution: *Goal Difficulty*, *Team Kaizen Experience*, *Work Area Routineness*, *Affective Commitment to Change* and *Internal Processes*. These variables were all significant at the

0.10/6 = 0.0167 level using the GEE "model based" standard error estimates, except for *Goal Difficulty*, which had a p-value of 0.0236. Thus, both of the models produced through backward selection procedure applied to the GEE estimates contained *Internal Processes*. The automated backward selection procedure using OLS produced the following solution: *Goal Difficulty*, *Team Autonomy*, *Team Functional Heterogeneity*, *Team Leader Experience*, *Internal Processes*, *Tool Appropriateness* and *Tool Quality*. This was the fifth best seven variable solution using the SAS RSQUARE selection procedure, the best seven variable using the MAXR selection procedure, and had the lowest C_p overall. However, using GEE "model based" standard error estimates, *Tool Appropriateness* had a p-value greater than 0.10 – i.e., much greater than 0.10/p. The automated stepwise selection procedure using OLS estimates produced the following solutions: *Team Autonomy*, *Team Kaizen Experience*, *Team Leader Experience* and *Internal Processes*. This was the best four variable solution using the SAS RSQUARE selection procedure and the MAXR selection procedure. However, *Team Kaizen Experience* had a p-value greater than 0.2 – i.e., much greater than 0.10/p – using GEE "model based" standard error estimates.

It should be noted that all of the models produced contain *Internal Processes* and three out of four contain *Team Autonomy* and/or *Goal Difficulty*. In addition, two models contain *Team Leader Experience*, *Team Kaizen Experience*, and/or *Tool Quality*. Due to the non-significant parameters using GEE "model based" standard error estimates for the two solutions produced through the OLS automated search procedures, the two competing solutions retained were those produced through the manual backward selection procedures. However, additional "good" subsets that were identified through the examination of the RSQUARE, ADJRSQ, CP and MAXR criteria were also as examined, as described more presently. The two initial competing solutions are show in Tables 23 and 24.

Table 23. Initial Regression Model for Task KSA (based on SE_{MB} $\hat{\beta}_{GEE}$)

	$\hat{\beta}_{GEE}$	SE_{MB} $\hat{\beta}_{GEE}$	SE_E $\hat{\beta}_{GEE}$	$\hat{\beta}_{OLS}$	SE $\hat{\beta}_{OLS}$
Intercept	.441	.504 (.382)	.758 (.561)	.385	.609 (.824)
Team Autonomy	.383	.084 (.000)	. 122 (.002)	.341	. 108 (.003)
Internal Processes	.500	.102 (.000)	.205 (.015)	.549	.127 (.000)
OLS R^2 *= .535; OLS* R_a^2 *= .515;* $F_{2,48}$ *= 27.59**** *GEE* R^2 *= .533; GEE* R_a^2 *= .513;* ρ *= 0.411*					

Table 24. Initial Regression Model for Task KSA (based on $SE_E\ \hat{\beta}_{GEE}$)

	$\hat{\beta}_{GEE}$	$SE_{MB}\ \hat{\beta}_{GEE}$	$SE_E\ \hat{\beta}_{GEE}$	$\hat{\beta}_{OLS}$	$SE\ \hat{\beta}_{OLS}$
Intercept	.080	.577 (.890)	.923 (.931)	.198	.698 (.778)
Goal Difficulty	.138	.061 (.024)	.021 (.000)	.115	.066 (.092)
Team Kaizen Experience	-.528	.070 (.000)	. 074 (.000)	-.425	.109 (.000)
Work Area Routineness	.112	.032 (.000)	.025 (.000)	.102	.052(.005)
Affective Commitment to Change	.291	.121 (.016)	.057 (.000)	.264	.106 (.045)
Internal Processes	.530	.130 (.000)	.157 (.001)	.543	.127 (.000)
*OLS R^2 = .635; OLS R_a^2 = .595; $F_{5,45}$ = 15.57**** *GEE R^2 =.626; GEE R_a^2 = .585; ρ = -0.071*					

Although several "good" subsets were tested – particularly the solutions produced through the MAXR procedure, since these all had relatively small p-values for predictors using OLS estimates – one "good subset" is particularly promising: *Goal Difficulty*, *Team Autonomy*, *Team Kaizen Experience*, *Team Leader Experience*, *Work Area Routineness*, *Affective Commitment to Change* and *Internal Processes*. This "good" subset was identified since it represented the best seven variable solution under the MAXR procedure and the third best three variable solution using the RSQUARE procedure. It was also the 10th best solution using the CP selection procedure, with $C_p \approx 9$. It should also be noted that this "good" subset combines the two solutions above. Three out of four of the p-values were less than 0.1/8 = 0.0125 and the remaining four p-values were all less than 0.05. Thus, although this represents an increased overall risk of family confidence, given the exploratory descriptive nature of this study and the purpose of identifying significant predictors, it appears reasonable to consider this solution, as an alternative to the two competing solutions above, which contain a smaller set of predictors. This solution is shown as Table 25.

Table 25. Final Regression Model for Task KSA

	$\widehat{\beta}_{GEE}$	$SE_{MB}\ \widehat{\beta}_{GEE}$	$SE_{E}\ \widehat{\beta}_{GEE}$	$\widehat{\beta}_{OLS}$	$SE\ \widehat{\beta}_{OLS}$
Intercept	.219	.543 (.687)	.725 (.763)	-.016	.635 (.981)
Goal Difficulty	.119	.119 (.032)	.029 (.000)	.098	.061 (.113)
Team Autonomy	.234	.234 (.014)	.114 (.040)	.273	.094 (.006)
Team Kaizen Experience	-.398	.095 (.000)	.055 (.000)	- .316	.108 (.006)
Team Leader Experience	-.195	.084 (.020)	.029 (.000)	-.168	.082 (.045)
Work Area Routineness	.094	.030 (.002)	.014 (.000)	.075	.032 (.026)
Affective Commitment to Change	.222	.112 (.049)	.079 (.005)	.170	.119 (.159)
Internal Processes	.465	.123 (.000)	.196 (.017)	.453	.120 (.000)
OLS $R^2 = .717$; *OLS* $R_a^2 = .671$; $F_{7,43} = 15.55$*** *GEE* $R^2 = .706$; *GEE* $R_a^2 = .658$; $\rho = -0.071$					

As a final step, residual analysis was performed on the final model to identify departures from normality, bias, influential observations, etc. There was no evidence of problems in the model. There were no apparent departures from linearity or significant evidence of departures from randomness in the residuals – i.e., Wald-Wolfowitz runs. Only two residuals out of 51 were more than two standard deviations from zero. However, both of these two cases were greater than 2.5 – one was equal to 2.58, the other was equal to 2.74. However, none of the Cook's distance values were significant ($p < 0.05$) or nearly significant ($p \leq 0.10$) for the final regression model and the maximum Cook's distance – which was for one of the two outlying observations – was 0.904 ($p = 0.519$).

4.3.5 Model of Impact on Area

The manual backward selection procedure using GEE "model based" standard error estimates, the automated backward and stepwise selection procedures using OLS standard error estimates, and the two hierarchical search procedures both converged upon the same two predictors: *Team Autonomy* and *Action Orientation*. This was also the best two variable solution under the RSQUARE selection procedure and the MAXR selection procedure. It was the best two variable solution and the fifth best solution overall using the CP selection procedure, although since $C_p < p$, some bias may be indicated. This solution is shown in Table 26. In general, the OLS and GEE "empirical" standard errors were very similar to, but slightly greater than, the "model based" standard errors, since the intraclass correlation coefficient was negative ($\rho = -0.071$) indicating more variation between versus within groups. However, all predictors are significant at the $\alpha = 0.05/p = 0.05/4 = 0.025$ level, so the substantive conclusions from all models are the same.

Table 26. Initial Regression Model for Impact on Area

	$\hat{\beta}_{GEE}$	SE_{MB} $\hat{\beta}_{GEE}$	SE_E $\hat{\beta}_{GEE}$	$\hat{\beta}_{OLS}$	SE $\hat{\beta}_{OLS}$
Intercept	1.384	.599 (.020)	.736 (.060)	1.497	.621 (.020)
Team Autonomy	.476	.136 (.000)	.198 (.016)	.486	.137 (.001)
Action Orientation	.290	.054 (.000)	.058 (.000)	.249	.077 (.002)
OLS $R^2 = .433$; *OLS* $R_a^2 = .409$; $F_{2,48} = 18.27$*** *GEE* $R^2 = .428$; *GEE* $R_a^2 = .404$; $\rho = -0.071$					

As described in the previous section, after identifying the initial solution, other "good" subsets from the automated search procedures were investigated to determine whether it would be possible to add additional predictors to the model. Most of these search results were not promising – i.e., resulted in individual p-values greater than 0.10 or multiple p-values greater than 0.05. However, one three variable solution is particularly promising. The regression model containing *Team Autonomy*, *Management Support* and *Action Orientation* was the best three variable solution using the RSQUARE and MAXR criteria and also the best solution overall using the CP criteria (i.e., the solution with the smallest C_p value, although since $C_p < p$, some bias may be indicated). In addition, using the GEE "model based" standard error estimates, all p-values are significant at the 0.05 level and all p-values except for *Management Support* are significant at the 0.05/p = 0.05/4 = 0.025 level. Thus, although *Management Support* is not significant at the 0.05/p = 0.05/4 = 0.025 level, the evidence appears strong enough ($p < 0.05$) to warrant its inclusion in the model. Thus, the three variable solution appears slightly preferred over the two variable solution. This solution is shown in Table 27. Note that in the OLS estimates, *Management Support* has a p-value less than 0.10 but greater than 0.05.

Table 27. Final Regression Model for Impact on Area

	$\hat{\beta}_{GEE}$	SE_{MB} $\hat{\beta}_{GEE}$	SE_E $\hat{\beta}_{GEE}$	$\hat{\beta}_{OLS}$	SE $\hat{\beta}_{OLS}$
Intercept	.953	.623 (.126)	.893 (.286)	.904	. 699 (.202)
Team Autonomy	.342	.149 (.022)	.198 (.016)	.330	.160 (.045)
Management Support	.262	.132 (.046)	.093 (.005)	.272	.156 (.087)
Action Orientation	.243	.058 (.000)	. 076 (.001)	. 256	.076 (.002)
OLS $R^2 = .467$; *OLS* $R_a^2 = .433$; $F_{3,47} = 13.72$*** *GEE* $R^2 = .466$; *GEE* $R_a^2 = .432$; $\rho = -0.071$					

As a final step, residual analysis was performed on the final model to identify departures from normality, bias, influential observations, etc. There was no evidence of problems in the model. There were no apparent departures from linearity or significant evidence of departures from randomness in the residuals – i.e., Wald-Wolfowitz runs. Only two residuals out of 51 were more than two standard deviations from zero. However, both of these two cases were close to 3.0 – one was slightly greater than 3.0 and one was slightly less than 3.0. However, none of the

Cook's distance values were significant ($p < 0.05$) or nearly significant ($p \leq 0.10$) for the final regression model and the maximum Cook's distance (which was actually for a non-outlying observation) was only 0.283 ($p = 0.888$).

Finally, the three-way and two-way interactions of the variables in the final model were tested to determine whether there was evidence of significant interactions of these terms. None of the interaction terms were significant.

4.3.6 Model of Overall Perceived Success

The manual backward selection procedure using GEE "model based" standard error estimates, the automated backward and stepwise selection procedures using OLS standard error estimates, and the two hierarchical search procedures both converged upon the same model with a single predictor: *Tool Quality*. The backward selection procedure using the GEE "empirical" standard error estimates suggested a two variable solution: *Team Leader Experience* and *Tool Quality*. However, *Team Leader Experience* was not significant using GEE model based standard error estimates ($p > 0.10$; therefore $p > 0.10/p = 0.10/3 = 0.033$). This model, was in fact, the next to last solution in the manual backward selection procedure – i.e., *Team Leader Experience* was removed to yield the final solution. Therefore, there is a lack of strong evidence that *Team Leader Experience* should be included in the model along with *Tool Quality*. Table 28 shows the coefficients and the standard error estimates for both the GEE and OLS models.

Table 28. Final Regression Model for Overall Perceived Success

	$\widehat{\beta}_{GEE}$	SE_{MB} $\widehat{\beta}_{GEE}$	SE_E $\widehat{\beta}_{GEE}$	$\widehat{\beta}_{OLS}$	SE $\widehat{\beta}_{OLS}$
Intercept	1.997	1.001 (.047)	.533 (.000)	2.152	1.054 (.047)
Tool Quality	.621	.215 (.004)	.109 (.000)	.595	.226 (.011)
OLS $R^2 = .124$; *OLS* $R_a^2 = .106$; $F_{1,49} = 6.96$** *GEE* $R^2 = .123$; *GEE* $R_a^2 = .105$; $\rho = -0.057$					

After identifying the initial solution, other "good" subsets from the automated search procedures were investigated. However, none of these subsets resulted in a full set of significant predictors and the R^2 values of all sets was low – i.e., less than 0.25. As a final step, residual analysis was performed on the final model to identify departures from normality, bias, influential, etc. There was no significant evidence of departures from randomness in the residuals – i.e., runs. There was some indication of departures from linearity due to the truncated distribution of the variable (to be discussed more in Chapter 5), although this departure did not appear extreme. Only three residuals out of 51 were more than two standard deviations from zero. However, all of these cases were greater than 2.5 and two of these cases were greater than 3.0 (but less than 4.0). This indicates that the model does not fit the data very well for these three cases. However, in these three cases, the response variable itself was an outlier. These three cases were the only three cases in the data set where *Overall Perceived Success* was less than 4.0. Thus, it appears that, for values of *Overall Perceived Success* less than 4.0 the model is not very accurate in prediction. However, for *Overall Perceived Success* values greater than 4.0, the model seems to be fairly accurate in prediction – i.e., there were no outliers for this range of response data. In addition, none of the Cook's distance values were significant ($p < 0.05$) or nearly significant ($p \leq 0.10$) for the final regression model and the maximum Cook's distance – which was for one of the outlying observations – was only 0.204 ($p = 0.816$).

4.3.7 Model of % of Goals Met

The manual backward selection procedure using GEE "model based" standard error estimates, the automated backward and stepwise selection procedures using OLS standard error estimates, and the two hierarchical search procedures both converged upon the same regression model with only one predictor: *Team Kaizen Experience*. In addition, in the next to last step in the backward selection procedure using GEE "model based" standard error estimates, *Event Planning Process* – i.e., number of hours spent planning the event – was significant at the 0.05 level but not the 0.10/3 = 0.033 level. Both models are shown in the following two tables (Table 29 and Table 30).

The GEE "empirical" standard error estimates suggested the following solution: *Goal Difficulty*, *Team Functional Heterogeneity*, *Team Kaizen Experience*, *Management Support* and *Event Planning Process*. However, using GEE "model based" standard error estimates, all p-values except for *Team Kaizen Experience* were greater than 0.05 and all p-values except for *Team Kaizen Experience* and *Event Planning Process* were greater than 0.10. Thus, this alternative solution lacks support and is not adopted.

Table 29. Initial Regression Model for % of Goals Met (based on $SE_{MB}\ \widehat{\beta}_{GEE}$)

	$\widehat{\beta}_{GEE}$	$SE_{MB}\ \widehat{\beta}_{GEE}$	$SE_{E}\ \widehat{\beta}_{GEE}$	$\widehat{\beta}_{OLS}$	$SE\ \widehat{\beta}_{OLS}$
Intercept	.236	.047 (.000)	.050 (.000)	.150	.080 (.068)
Team Kaizen Experience	-.471	.074 (.000)	.087 (.000)	-.380	. 106 (.001)
*OLS R^2 = . 207; OLS R_a^2 = . 191; $F_{1,47}$ = 12.77*** GEE R^2 = .188; GEE R_a^2 =.171; ρ = -0.067*					

Table 30. Final Regression Model for % of Goals Met (based on $SE_{E}\ \widehat{\beta}_{GEE}$)

	$\widehat{\beta}_{GEE}$	$SE_{MB}\ \widehat{\beta}_{GEE}$	$SE_{E}\ \widehat{\beta}_{GEE}$	$\widehat{\beta}_{OLS}$	$SE\ \widehat{\beta}_{OLS}$
Intercept	.125	.077 (.020)	.077 (.106)	.045	.103 (.662)
Team Kaizen Experience	-.486	.067 (.000)	.057 (.000)	.-.392	.105 (.001)
Event Planning Process	.149	.075 (.047)	.074 (.044)	.136	.084 (.114)
*OLS R^2 = . 247; OLS R_a^2 = .216; $F_{2,48}$ = 7.89** GEE R^2 = .226; GEE R_a^2 =.194; ρ = -0.071*					

Additional "good subsets" were investigated, but no promising candidates were found. As a final step, residual analysis was performed on the final model to identify departures from normality, bias, influential, etc. There was no significant evidence of departures from randomness in the residuals – i.e., runs. There was some indication of departures from linearity due to the skewed and truncated distribution of the variable (to be discussed more presently), although this departure did not appear extreme. Only two residuals out of 51 were more than two standard deviations from zero. One of these residuals was only slightly greater than 2.0 (2.15). However, the second residual was greater than 5.0 (5.32). This case was the only event in the data set that did not achieve any of its goals – i.e., % of Goals met was zero. Thus, the level of the response variable itself was an outlier. None of the Cook's distance values were significant ($p < 0.05$). However, the Cook's distance value was nearly significant ($p \leq 0.10$) for the case that was 5.0 standard deviations from zero. This suggests that there is some evidence that this unusual case is exerting a relatively large influence on the model. There is no evidence of a coding mistake or respondent error for this observation, so there is no justification for removing it from the data set (Field, 2005). However, it should be noted that the regression results could have been somewhat different had this case not existed.

In general, the explanatory power of these models is low. It seems likely that this is due, at least in part, to the fact that, while the response variable *% of Goals Met* is continuous, the distribution of the response variable was extremely skewed and truncated. In the data set of 51 teams, 35 teams met all their goals and therefore had the same response value – i.e., 1.0 prior to the log transformation of the variable and 0.0 after the transformation. Therefore, to investigate whether any additional variables contribute to the prediction of the relative goal achievement, *% of Goals Met* was transformed into an ordinal, binomial variable – i.e., 1.0 if all goals were met, 0 otherwise – and a logistic regression was performed using SAS PROC GENMOD and the GEE "model based" estimates. The models were executed using a logit link function and a binomial distribution. In addition, the "descending" option was specified such that the results model the probability of a response equal to 1.0 and thus are consistent with the signs in Gaussian regression model above. In addition, a logistic regression model without the GEE correction was created using the "Logistic Regression" procedure in SPSS, which uses weighted least squares (WLS). Again, a backward selection process based on the likelihood ratio statistic (as recommended by Field, 2005) was used.

Generally, it is not advisable to transform a continuous variable into an ordinal variable, due to loss in sensitivity – i.e., treating all events with *% of Goals Met* less than 100% as having the same relative level of success. However, in this case, the analysis modeling *% of Goals Met* as a continuous variable had already been performed and was shown to be low in explanatory power. The logistic regression answers a similar, but related question: *What variables differentiate events that met 100% of the goals from events that did not?* Again, backward selection processes using both GEE "model-based" standard error estimates and GEE "empirical" standard error estimates were performed. Both selection procedures converged upon the same solutions. Only one model was located where all predictors were less than 0.10/p. However, a second solution was located at an earlier stage in the selection process where all p were less than 0.05. These two "good" solutions are shown in Tables 31 and 32. In addition to the two backwards procedures, the two hierarchical regression procedures were also employed. The first – i.e., regressing process variables then input variables – failed to converge upon a solution. The second – i.e., regressing all variables except for *Action Orientation* – yielded a reduced form of the final solution: *Team Kaizen Experience*, *Team Leader Experience* and *Event Planning Process*. However, the logistic regression performed using SPSS without GEE correction converged upon the same five factor solution that was ultimately adopted (Table 32).

Table 31. Initial Logistic Regression Model for % of Goals Met

	$\widehat{\beta}_{GEE}$	SE_{MB} $\widehat{\beta}_{GEE}$	SE_E $\widehat{\beta}_{GEE}$	$\widehat{\beta}_{WLS}$	SE $\widehat{\beta}_{WLS}$
Intercept	1.940	.851 (.023)	.694 (.005)	1.571	1.00 (.119)
Team Kaizen Experience	-2.958	1.073 (.006)	1.174(.012)	-2.181	1.119(.051)
Team Leader Experience	-2.243	.908 (.014)	.404 (.000)	-1.875	.904 (.038)
Event Planning Process	2.283	.833 (.006)	.956 (.017)	2.046	.986 (.038)
*WLS R^2 = . 395 (Nagelkerke);WLS χ^2 (3)= 16.83*** *GEE ρ = -0.071*					

Table 32. Final Logistic Regression Model for % of Goals Met

	$\widehat{\beta}_{GEE}$	SE_{MB} $\widehat{\beta}_{GEE}$	SE_E $\widehat{\beta}_{GEE}$	$\widehat{\beta}_{WLS}$	SE $\widehat{\beta}_{WLS}$
Intercept	17.383	6.677 (.009)	5.954 (.004)	16.388	6.34 (.010)
Goal Difficulty	-1.979	.978 (.043)	.942 (.036)	-1.876	.921 (.042)
Team Kaizen Experience	-4.254	1.314 (.001)	1.155 (.000)	-3.368	1.318(.011)
Team Leader Experience	-4.262	1.429 (.003)	828 (.000)	-3.723	1.412(.008)
Event Planning Process	4.554	1.435 (.002)	1.330 (.001)	4.161	1.520(.006)
Action Orientation	-1.901	.867 (.028)	.601 (.002)	-1.820	.860 (.034)
*WLS R^2 = .553 (Nagelkerke); WLS χ^2 (5)= 25.51**** *GEE ρ = -0.071*					

As can be seen from Tables 29 through 32, there is some similarity to the solution that modeled *% of Goals Met* as a continuous variable. Again, *Team Kaizen Experience* and *Event Planning Process* (hours planning) are significant predictors, and the direction of the signs are the same. In addition, the logistic regression models identified three additional predictors that differentiate more (100%) versus less (less than 100%) successful events: *Goal Difficulty*, *Team Leader Experience* and *Action Orientation*. Since, in the five variable solution – i.e., *Goal Difficulty*, *Team Kaizen Experience*, *Team Leader Experience*, *Event Planning Process* and *Action Orientation* – all variables were significant at the 0.05 level and all but two variables were significant at the 0.10/p = 0.10/5 = 0.02 level, the five variable solution appears to be slightly preferred over the three variable solution. It should be noted that this five variable solution is also quite similar to the solution implied through the "empirical" standard error estimates in the Gaussian regression. The interpretation of these results is discussed in Chapter 5.

As always, the final model was further tested for "goodness-of-fit" through residual analysis. There was no significant evidence of departures from randomness in the residuals – i.e., runs. It addition, the predicted versus observed classifications of cases was observed. In SPSS model without the GEE correction, using a cutoff probability of 0.5 (the default in SPSS), nine total cases were misclassified. Four events that met 100% of their goals (i.e., had a goal achievement value of 1) were incorrectly classified as 0's. Five events that met less than 100% of their goals (i.e., had a goal achievement value of 0) were incorrectly classified as 1's. However, when a plot of the observed groups and predicted probabilities was examined, it was noted that all four of the 1's that had been misclassified had a predicted probability very close to the cutoff probability – i.e., greater than 0.45. Thus, shifting the cutoff value slightly downward would have resulted in no misclassification for events with a goal achievement score of 1. In addition, this shift would not have caused any additional events with a goal achievement score of 0 to be misclassified. The five events with a goal achievement score of 0 that were misclassified as 1's all had actual *% of Goals Met* values greater than 50%. However, one of the events had a *% of Goals Met* value that was only slightly greater than 50% (51.1%). The other four events had a *% of Goals Met* value greater than 75% (the minimum value was 77.6%). A similar breaking point was observed using the GEE regression results – by lowering the cutoff probability somewhat, some of the misclassified 1's would have been correctly classified, without resulting in additional errors in the classification of zeros. However, one event with a goal achievement value of 1 would have remained misclassified using GEE model results – i.e., the cutoff probability could not be lowered far enough to include this event in group 1 without introducing classification errors in group 0. Thus, overall it seems that the model is fairly successful at predicting group membership for events that achieved 100% of their goals, but less successful at predicting group membership for events that did not achieve 100% of their goals. However, roughly two thirds of these events – 11 out of 16 using SPSS results and 10 out of 16 using GEE results – were still correctly classified.

Only two residuals out of 51 were more than two standard deviations from zero. One of these residuals was only slightly greater than 2.0 (2.11). However, the second residual was greater than 5.0 (5.58). These were two different cases than the outliers for the continuous distribution – i.e., Gaussian -- model. Although this case was classified as a 0 due to failure to meet 100% of the event goals, this team met 95% of the goals. This was highest level of *% of Goals Met* for all events that were less than 100% successful. Thus, it makes sense that this event may actually be more similar – i.e., share more common characteristics – with more versus less successful characteristics, and therefore have been misclassified as a 1. None of the Cook's distance values were significant ($p < 0.05$) or nearly significant ($p \leq 0.10$).

4.3.8 Summary of Final Regression Models

Table 33 summarizes the significant direct predictors for each outcome variable.

Table 33. Significant Direct Predictors of Outcome Variables

Outcome Variable	Significant Predictors (p-values)
Attitude	• Team Functional Heterogeneity (0.012) • Management Support (0.013) • Internal Processes (0.000)
Task KSA	• Goal Difficulty (0.032) • Team Autonomy (0.014) • Team Kaizen Experience (0.000) • Team Leader Experience (0.020) • Work Area Routineness (0.002) • Affective Commitment to Change (0.049) • Internal Processes (0.000)
Impact on Area	• Team Autonomy (0.022) • Management Support (0.046) • Action Orientation (0.000)
Overall Perceived Success	• Tool Quality (0.004)
% of Goals Met	• Goal Difficulty (n.s. cont., dicht.) • Total Kaizen Experience (.000 cont., .000 dicht.) • Team Leader Experience (n.s. cont., dicht.) • Event Planning Process (.047 cont., dicht.) • Action Orientation (n.s. cont., dicht.)

As a final check, the VIF were examined for the final models from the exploratory regression (see Table 33), even though the initial analysis of VIF had indicated that multicollinearity was not severe. This was done to verify that the final set of predictor variables in each model were not strongly related, while the initial VIF analysis considered relations between all fourteen potential predictor variables. Table 34 presents the VIF for the final models. In all the models, the maximum observed VIF was 1.75 and the average VIF was less than 1.5 – i.e., fairly close to 1.00 and substantially less than 3.0.

Table 34. VIF Values for Final Regression Models

Outcome	Predictor	VIF_{OLS}	VIF_{GEE}
Attitude	Team Functional Heterogeneity	1.02	1.01
	Management Support	1.19	1.19
	Internal Processes	1.18	1.17
	Average VIF	1.13	1.13
	Max VIF	1.19	1.19
Task KSA	Goal Difficulty	1.16	1.13
	Team Autonomy	1.47	1.42
	Team Kaizen Experience	1.28	1.11
	Team Leader Experience	1.22	1.19
	Work Area Routineness	1.17	1.11
	Affective Commitment to Change	1.74	1.69
	Internal Processes	1.72	1.69
	Average VIF	1.34	1.33
	Max VIF	1.74	1.69
Impact on Area	Team Autonomy	1.66	1.62
	Management Support	1.48	1.42
	Action Orientation	1.17	1.11
	Average VIF	1.43	1.38
	Max VIF	1.66	1.62
Overall Perceived Success	Tool Quality	1.00	1.00
	Average VIF	1.00	1.00
	Max VIF	1.00	1.00
% of Goals Met	Goal Difficulty	1.31	1.29
	Total Kaizen Experience	1.26	1.12
	Team Leader Experience	1.40	1.35
	Event Planning Process	1.30	1.29
	Action Orientation	1.24	1.20
	Average VIF	1.30	1.25
	Max VIF	1.40	1.35

4.4 Mediation Analysis to Test H9 & H10

Mediation occurs when one variable acts indirectly upon a second variable through a third, mediating variable (Baron & Kenny, 1986; Kenny, 2006). The mediation effect can be partial – i.e., the first variable has a significant indirect effect on the second variable, as well as a significant direct effect – or full – i.e., the first variable only has a significant effect through the third variable.

Mediation analysis is based upon a hypothesized causal model – i.e., a hypothesized set of causal relationships between several variables: an input variable, an intervening or process variable, and an outcome (MacKinnon et al., 2000; Kenny, 2006). Thus, mediation analysis results are only valid if the direction of causality specified in the hypothesized causal model is correct. In observation studies, such as this one, it is generally impossible to completely rule out the possibility that the causal relationship is actually the reverse of that specified or non-existent

– i.e., that both outcomes and predictors are correlated with the true, unknown causes of the outcomes but that the predictor does not determine the level of the cause. (See the extended discussion of causality in observational – particularly cross-sectional – studies in Chapter 5).

However, mediation analysis is useful for testing theory, since, if a hypothesized mediation relationship is not demonstrated, this lends support to the conclusion that the hypothesized model is not correct (MacKinnon et al., 2000). If the hypothesized mediation relationship is demonstrated, this indicates that the proposed causal model could be correct and warrants further investigation – preferably through additional research designs involving experimental control. See MacKinnon, Krull and Lockwood (2000) for further discussion of the benefits of following up initial cross-sectional mediation analyses with replications, preferably involving experimental control.

Mediation is often explained using the two equations (e.g., see Alwin & Hauser, 1975, and Kenny, 2006 for more descriptions of mediation). Equation 6 depicts a direct, unmediated relationship between X and Y, where c is the strength of the path. Equation 7 depicts a situation where Z at least partially mediates the relationship between X and Y, through paths a and b. X may also still have a direct effect on Y, if there exists an additional path c' (not shown in Equation 7below) which is significant after controlling for the effect of Z.

$$X \overset{c}{\rightarrow} Y \tag{6}$$

$$X \overset{a}{\rightarrow} Z \overset{b}{\rightarrow} Y \tag{7}$$

In the current research, event process factors – i.e., *Affective Commitment to Change*, *Action Orientation*, *Internal Processes*, *Tool Appropriateness* and *Tool Quality* – are hypothesized to serve as at least partial mediators between event input factors and event outcomes factors. This is expressed in H9 and H10 (see also Figure 1):

- H9: Event process factors will partially mediate the relationship of event input factors and social system outcomes.
- H10: Event process factors will partially mediate the relationship of event input factors and technical system outcomes.

Judd and Kenny (1981) and Baron and Kenny (1986) present a four step method of analyzing mediation relationships that includes the testing of all four paths – a, b, c and c.' However, Kenny (2006) and Kenny, Kashy, and Bolger (1998) indicate that there is a more parsimonious method of testing for a mediation effect. Only paths a and b must be tested to demonstrate mediation. The steps for testing these paths are as follows:

1. Regress the mediator variable (Z) on the initial variable (X). This tests whether there is a significant relationship between X and Z – i.e., whether path a in Equation 7 is significant. In this research, this test is accomplished by performing simple linear regression using GEE.
2. Next, regress both Z and X on Y. This will indicate whether Z has a significant effect on Y while controlling for X – i.e.,, whether path b in Equation 7 is significant. In this research, this test is accomplished by performing multiple linear regression using GEE. If both path a and path b are significant, the indirect effect of X on Y can then be calculated as a * b (MacKinnon et al., 1995).
3. Optional: Path c' can also be tested using the output of step 2. If path c' is significant and paths a and b are significant this is consistent with a partial mediation hypothesis. If path c' is not significant and paths a and b are significant, this is consistent with a full mediation hypothesis.

Although mediation analysis was first performed using OLS regression, it can be performed with other types of regression analysis, such as logistic regression, structural equations modeling and multilevel modeling – i.e., HLM or "mixed models" (Kenny, 2006). Although it is not clear whether anyone to date has performed mediation analysis using GEE, there appears to be no theoretical reason why mediation analysis cannot be performed using GEE estimates versus OLS or logistic regression estimates, since both are asymptotically unbiased estimates of the same unknown variables.

As described above, one way to demonstrate significance of the overall hypothesized mediation effect is to demonstrate the significance of each path being tested – i.e., path a and path b. However, to preserve the overall Type I error protection of the test, the alpha level must be adjusted for the number of simultaneous tests – i.e., the number of parameters being tested to test each hypothesized mediation relationship. If the desire is only to test whether the mediation hypothesis holds, and not to differentiate between a full and partial mediation hypothesis – i.e., to estimate only path a and path b – a 0.05/2 = 0.025 significance threshold for individual tests will preserve a 0.05 family confidence level. If the intent is to also test whether the hypothesized mediation effect is consistent with a partial or full hypothesis -- i.e., to estimate path c' – a 0.05/3 = 0.0167 confidence level is necessary to preserve a 0.05 family confidence level. Since in the current research it is of interest to test whether the mediation is partial or full, an alpha value of 0.0167 will be used.

In addition to the separate tests of regression parameters, Sobel (1982) proposed a combined test which is widely used. However, the Sobel test requires that the standard errors of paths a and b are independent, and this assumption appears questionable in the current research due to the hypothesized correlated error structure and the fact that a common working correlation matrix is used in estimating the regression parameters. Thus, the approach of separately testing each path with a correction for overall family confidence is used.

In the current research, there are five event process variables – i.e., *Affective Commitment to Change*, *Action Orientation*, *Internal Processes*, *Tool Appropriateness* and *Tool Quality* – and nine event input factors. However, for each outcome measure, only the event process variables shown to have a significant relationship to the outcome will be tested for mediation effects. The next five sections describe the tests of mediation effects of the five outcomes variables. Following each mediation analysis, for any significant mediation effects, the sign, but not the significance, of the regression coefficient when the relevant outcome variable was regressed on the input variable only was examined in order to check for suppressor effects. A suppressor effect occurs when the directionality of the input variable's relationship to the outcome measure is different for the direct effect versus the mediated, indirect effect (MacKinnon et al., 2000).

4.4.1 Mediation Analysis for Attitude

As Table 33 demonstrates, there was only one event process factor that was a significant predictor for *Attitude* – *Internal Processes*. Thus, in the first step of the mediation analysis, *Internal Processes* is separately regressed on all nine event input factors to determine which event input factors have a significant relationship to *Internal Processes*. The results of these separate regressions are shown in Table 35. Note, parameter estimates and standard error estimates are only shown for significant regressions. The standard error estimates are the GEE "model based" standard estimates and, as in the regression models used to test the other study hypotheses, an exchangeable correlation matrix, Gaussian distribution and identity link function are used.

Table 35. Internal Processes on Input Variable (X) Regressions (path a)

Input Variable (X)	Coefficient (a)	Std. Error	p-value
Goal Clarity	.5706	.0714	.0000
Goal Difficulty			.7342
Team Autonomy	.4076	.1057	.0001
Team Functional Heterogeneity			.7745
Team Kaizen Experience			.1347
Team Leader Experience			.9767
Management Support	.3363	.1156	.0036
Event Planning Process			.0759
Work Area Routineness			.9892

For regressions where the coefficient – i.e., path a – was significant at the $\alpha = 0.05/3 = 0.0167$ level – i.e., *Goal Clarity*, *Team Autonomy* and *Management Support* – a second regression is performed where *Attitude* is regressed on both the input variable (X) and *Internal Processes*. These results are shown in Table 36.

Table 36. Attitude on Internal Processes and Input Variable (X) Regressions (path b and path c')

Input Variable (X)	Coefficient (c')	Std. Error	p-value	Internal Processes Coefficient (b)	Internal Processes Std. Error	p-value
Goal Clarity	.1474	.1354	.2762	.7251	.1394	.0000
Team Autonomy	.1847	.1119	.0988	.7005	.1330	.0000
Management Support	.2848	.1052	.0068	.6785	.1207	.0000

Path b – i.e., the impact of *Internal Processes* on *Attitude* while controlling for the predictor (X) – was significant for all three regressions at the $\alpha = 0.05/3 = 0.0167$ level. Thus, mediation analysis results for *Goal Clarity*, *Team Autonomy* and *Management Support* are consistent with the mediation hypothesis – i.e., the hypothesis that these variables impact *Attitude* indirectly through *Internal Processes*. In addition, path c' was significant for *Management Support*, which is consistent with a partial mediation effect – i.e., the hypothesis that *Management Support* has both a direct and an apparent indirect effect on *Attitude*, which agrees with that the fact that *Management Support*, along with *Internal Processes* was one of the three variables in the final regression model. Path c' was non-significant for *Goal Clarity* and *Team Autonomy*, which is consistent with a full mediation effect – i.e., the hypothesis that *Goal Clarity* and *Team Autonomy* significantly affect *Attitude*, but only indirectly through *Internal Processes*.

Following the mediation analysis, a post-hoc analysis was performed to further test the robustness of the mediation hypotheses. In this analysis, *Internal Processes* was regressed simultaneously on *Goal Clarity*, *Team Autonomy* and *Management Support*. As shown in Table 37, only the regression coefficient for *Goal Clarity* was clearly significant in this regression ($p < 0.05$). Thus, while the relationships between *Team Autonomy*, *Internal*

Processes and *Attitude*, and *Management Support*, *Internal Processes* and *Attitude* are consistent with the mediation hypothesis, there is weaker support for these relationships than for the relationship between *Goal Clarity, Internal Processes* and *Attitude*. Specifically, after controlling for *Goal Clarity, Team Autonomy* and *Management Support* are no longer significant predictors of *Internal Processes*.

Table 37. Internal Processes on Goal Clarity, Team Autonomy and Management Support

Input Variable (X)	Coefficient (a)	Std. Error	p-value
Goal Clarity	.4448	.1021	.0000
Team Autonomy	.1543	.1252	.2176
Management Support	.0956	.1133	.3984

Table 38 summarizes the findings of the mediation analysis and the total strength of path for the mediation effects – i.e., a*b. Note that the effects of *Team Autonomy* and *Management Support* are italicized, since a post hoc analysis indicated that they are not significant unique predictors of *Internal Processes* after controlling for *Goal Clarity*. Only a point estimate is provided since calculating the confidence interval would require assumption that the standard errors of a and b are independent. Note that the calculated mediated effect is the raw, unmoderated mediated effect. The actual mediated effect will likely be slightly smaller, since, in most cases, the coefficient of *Internal Processes* in Table 36 is slightly greater than the coefficient of *Internal Processes* in the final regression model ($\widehat{\beta}_{GEE}$ = 0.694, see Table 22), due to the moderation of *Internal Processes* by *Management Support* and *Team Functional Heterogeneity*. Implications of these findings are discussed in Chapter 5.

Table 38. Summary of Mediation Analysis Results for Attitude

Input Variable (X)	Mediator Variable	Total mediated effect (a*b)	Partial or Full
Goal Clarity	Internal Processes	.4137	Full
Team Autonomy	*Internal Processes*	*.2855*	*Full*
Management Support	*Internal Processes*	*.2282*	*Partial*

4.4.2 Mediation Analysis for Task KSA

There were two event process factors that were significant predictors of *Task KSA –Internal Processes* and *Affective Commitment to Change*. For *Internal Processes* the output from Table 35 can be used to proceed directly to the second step in the mediation analysis (since the path a calculations remain the same). For regressions where the coefficient – i.e., path a – was significant at the $\alpha = 0.05/3 = 0.0167$ level – i.e., *Goal Clarity*, *Team Autonomy* and *Management Support* – a second regression is performed where *Task KSA* is regressed on both X and *Internal Processes*. These results are shown in Table 39.

Table 39. Task KSA on Internal Processes and Input Variable (X) Regressions (path b and path c')

Input Variable (X)	Coefficient (c')	Std. Error	p-value	Internal Processes Coefficient (b)	Internal Processes Std. Error	p-value
Goal Clarity	.2273	.1285	.0770	.6104	.1214	.0000
Team Autonomy	.3825	.0843	.0000	.4996	.1019	.0000
Management Support	.2495	.1017	.0142	.6143	.1114	.0000

These results suggest that path b – i.e., the impact of *Internal Processes* on *Task KSA* while controlling for X – is significant for all three input variables at the $\alpha = 0.05/3 = 0.0167$ level In addition, path c' was significant at the 0.0167 level for *Team Autonomy* and *Management Support*, which is consistent with a partial mediation effect. Path c' was non-significant for *Goal Clarity*, indicating a full mediation effect. For *Team Autonomy*, the apparent partial mediation effect is consistent with the finding that *Team Autonomy* is a significant predictor in the final regression model of *Task KSA*. The fact that *Management Support* was not significant in the final regression suggests that another variable in the final regression moderates the impact of *Management Support* on *Task KSA*. However, as mentioned previously, a post-hoc analysis revealed that, after controlling for *Goal Clarity*, *Team Autonomy* and *Management Support* are no longer significant predictors of *Internal Processes*. Thus, there is less support for the mediation hypotheses involving *Team Autonomy* and *Management Support* than for the mediation hypothesis involving *Goal Clarity*.

For *Affective Commitment to Change*, in the first step of the mediation analysis, *Affective Commitment to Change* is separately regressed on all nine event input factors to determine which event input factors have a significant relationship to *Affective Commitment to Change*. The results of these separate regressions are shown in Table 40. Note, parameter estimates and standard error estimates are only shown for significant regressions.

Table 40. Affective Commitment to Change on Input Variable (X) Regressions (path a)

Input Variable (X)	Coefficient (a)	Std. Error	p-value
Goal Clarity	0.6629	0.0627	0.0000
Goal Difficulty			0.0556
Team Autonomy	0.3497	0.1087	0.0013
Team Functional Heterogeneity			0.3913
Team Kaizen Experience			0.7641
Team Leader Experience			0.9646
Management Support	0.3378	0.1184	0.0043
Event Planning Process			0.3377
Work Area Routineness			0.4917

For regressions where the coefficient – i.e., path a – was significant at the α = 0.05/3 = 0.0167 level – i.e., *Goal Clarity*, *Team Autonomy* and *Management Support* -- a second regression is performed where *Task KSA* is regressed on both the input variable (X) and *Internal Processes*. These results are shown in Table 41.

Table 41. Task KSA on Affective Commitment to Change and Input Variable (X) Regressions (path b and path c')

Input Variable (X)	Coefficient (c')	Std. Error	p-value	Affective Commitment to Change Coefficient (b)	Affective Commitment to Change Std. Error	p-value
Goal Clarity	0.3565	0.1756	.0424	0.2543	0.1687	.1317
Team Autonomy	0.4954	0.0958	.0000	0.2298	0.1167	.0489
Management Support	0.3383	0.1228	.0059	0.3339	0.1358	.0139

Path b – i.e., the impact of *Internal Processes* on *Task KSA* while controlling for the predictor (X) – was significant only for *Management Support* at the α = 0.05/3 = 0.0167 level. Thus, the mediation analysis results for *Goal Clarity* and *Team Autonomy* do not appear consistent with the mediation hypothesis – i.e., the hypothesis that these variables impact *Task KSA* indirectly through *Affective Commitment to Change*. However, it should be noted that the p-value for the effect involving *Team Autonomy* was fairly low, providing some weaker support for a mediation effect. In addition, path c' was significant for *Management Support*, which is consistent with a partial mediation effect. However, as mentioned previously, the fact that *Management Support* was not significant in the final regression indicates that another variable in the final regression moderates the impact of *Management Support* on *Task KSA*. In addition, support for this model is especially tentative methodologically given the fact that, in this research, *Management Support* was measured antecedent to *Affective Commitment to Change*. Support for a mediation model involving *Management Support* and *Affective Commitment to Change* is valid only if it can be inferred that, as intended, *Management Support* measures a global variable rather than only the specific aspects of *Management Support* directly inquired about in the constituent items in the final survey scale.

Similarly to that performed for *Internal Processes*, following the mediation analysis, a post-hoc analysis was performed to further test the robustness of the mediation hypothesis. In this analysis, *Affective Commitment to Change* was regressed simultaneously on *Goal Clarity*, *Team Autonomy* and *Management Support*. As shown in Table 42, only the regression coefficient for *Goal Clarity* was clearly significant in this regression ($p < 0.05$). Thus, while the relationships between *Management Support*, *Affective Commitment to Change* and *Task KSA* are consistent with the mediation hypothesis, support for the mediation hypothesis is weakened by the fact that, after controlling for *Goal Clarity, Management Support* is no longer a significant, unique predictors of *Affective Commitment to Change*.

Table 42. Affective Commitment to Change on Goal Clarity, Team Autonomy and Management Support

Input Variable (X)	**Coefficient (a)**	**Std. Error**	**p-value**
Goal Clarity	.5993	.1196	.0000
Team Autonomy	.0319	.1157	.7827
Management Support	.0843	.1111	.4483

Table 43 summarizes the findings of the mediation analysis and the total strength of path for the mediated effects – i.e., a*b. Note that the effects of *Team Autonomy* and *Management Support* are italicized, since a post hoc analysis indicated that they are not significant unique predictors of *Internal Processes* and *Affective Commitment to Change* after controlling for *Goal Clarity*. Implications of the findings are discussed in Chapter 5.

Table 43. Summary of Mediation Analysis Results for Task KSA

Input Variable (X)	**Mediator Variable**	**Total mediated effect (a*b)**	**Partial or Full**
Goal Clarity	Internal Processes	.3155	Full
Team Autonomy	*Internal Processes*	*.2058*	*Partial*
Management Support	*Internal Processes*	*.2066*	*Partial*
Management Support	*Affective Commitment to Change*	*.1128*	*Partial*

4.4.3 Mediation Analysis for Impact on Area

There was one event process factor that was a significant predictor of *Impact on Area – Action Orientation*. Table 44 shows the results for regressing *Action Orientation* on all nine event input variables. The regressions were significant at the 0.0167 level for three variables – *Goal Difficulty*, *Team Autonomy* and *Work Area Routineness*.

Table 44. Action Orientation on Input Variable (X) Regressions (path a)

Input Variable (X)	**Coefficient (a)**	**Std. Error**	**p-value**
Goal Clarity			.1608
Goal Difficulty	-.4517	.1632	.0056
Team Autonomy	.6265	.2078	.0026
Team Functional Heterogeneity			.1573
Team Kaizen Experience			.4407
Team Leader Experience			.3531
Management Support			.6362
Event Planning Process			.5286
Work Area Routineness	0.2643	0.0835	.0016

Table 45 shows the regression of *Impact on Area* on both *Action Orientation* and the input variable (X) for the regressions where path a was significant.

Table 45. Impact on Area on Internal Processes and Input Variable (X) Regressions (path b and path c')

Input Variable (X)	Coefficient (c')	Std. Error	p-value	Action Orientation Coefficient (b)	Action Orientation Std. Error	p-value
Goal Difficulty	-.0666	.0897	.4581	.3768	.0539	.0000
Team Autonomy	.4757	.1363	.0005	.2905	.0541	.0000
Work Area Routineness	-.0352	.0621	.5711	.4013	.0714	.0000

These results indicate that path b – i.e., the impact of *Action Orientation* on *Impact on Area* while controlling for X – is significant for all three variables at the $\alpha = 0.05/3 = 0.0167$ level -- which is consistent with the mediation hypothesis. Path c' was also significant at the 0.0167 level for *Team Autonomy*, which is consistent with a partial mediation effect, which also agrees with the fact that *Team Autonomy*, along with *Action Orientation*, was one of the three variables in the final regression model. Path c' was non-significant for *Goal Difficulty* and *Work Area Routineness*, which is consistent with a full mediation effect for these two variables – i.e., under the full mediation hypothesis, *Goal Difficulty* and *Work Area Routineness* appear to significantly affect *Impact on Area*, but only indirectly through *Action Orientation*.

Following the mediation analysis, a post-hoc analysis was performed to further test the robustness of the mediation hypothesis. In this analysis, *Action Orientation* was regressed simultaneously on *Goal Difficulty*, *Team Autonomy* and *Work Area Routineness*. As shown in Table 46, this post hoc analysis indicated that all three variables were significant unique predictors of Action Orientation at the $\alpha = 0.05$ level.

Table 46. Action Orientation on Goal Difficulty, Team Autonomy and Work Area Routineness

Input Variable (X)	Coefficient (a)	Std. Error	p-value
Goal Difficulty	-.3493	.1196	.0210
Team Autonomy	.5028	.1157	.0086
Work Area Routineness	.1743	.1111	.0264

Table 47 summarizes the findings of the mediation analysis and the total strength of path for the mediation effects – i.e., a*b. As shown, in the mediation model, *Goal Difficulty* has a negative effect on *Impact on Area*, while *Team Autonomy* and *Work Area Routineness* have positive effects on *Impact on Area*. Further implications of all these relations will be discussed in Chapter 5.

Table 47. Summary of Mediation Analysis Results for Impact on Area

Input Variable (X)	Mediator Variable	Total mediated effect (a*b)	Partial or Full
Goal Difficulty	Action Orientation	-.1702	Full
Team Autonomy	Action Orientation	.1820	Partial
Work Area Routineness	Action Orientation	.1061	Full

4.4.4 Mediation Analysis for Overall Perceived Success

There was one event process factor that was a significant predictor for *Overall Perceived Success*: *Tool Quality*. Table 48 shows the results for regressing *Tool Quality* on all nine event input variables. The regressions were significant at the 0.0167 level for two variables – *Goal Clarity* and *Management Support*. In addition, one variable, *Team Autonomy*, had a small but non-significant p-value ($p < 0.05$).

Table 48. Tool Quality on Input Variable (X) Regressions (path a)

Input Variable (X)	Coefficient (a)	Std. Error	p-value
Goal Clarity	0.5448	0.2161	0.0117
Goal Difficulty			0.2350
Team Autonomy			0.0496
Team Functional Heterogeneity			0.7610
Team Kaizen Experience			0.1392
Team Leader Experience			0.6897
Management Support	0.5615	0.2056	0.0063
Event Planning Process			0.9833
Work Area Routineness			0.7480

Table 49 shows the regression of *Impact on Area* on both *Tool Quality* and the input variable (X) for the regressions where path a was significant.

Table 49. Impact on Area on Internal Processes and Input Variable (X) Regressions (path b and path c')

Input Variable (X)	Coefficient (c')	Std. Error	p-value	Action Orientation Coefficient (b)	Action Orientation Std. Error	p-value
Goal Clarity	.5184	.3887	.1824	.4989	.2350	.0338
Management Support	.2002	.3588	.5769	.5630	.2399	.0190

These results indicate that the path b – i.e., the impact of *Tool Quality* on *Overall Perceived Success* while controlling for the input variable (X) – is not significant for any of the regressions, providing a lack of clear support for the mediation hypothesis. However, in both regressions the p-value for the effect of *Tool Quality* on *Overall Perceived Success* was small ($p < 0.05$). In addition, for the regression of *Tool Quality* and *Management Support* on *Overall Perceived Success*, the p-value of path b was only slightly greater than 0.0167 (i.e., 0.0190). This suggests that, although not formally significant, this effect should likely be considered in the research results. In other words, there is relatively strong evidence that impact of *Management Support* on *Overall Perceived Success* is consistent with a full mediation effect.

Following the mediation analysis, a post-hoc analysis was performed to further test the robustness of the mediation hypothesis. In this analysis, *Tool Quality* was regressed simultaneously on *Goal Clarity* and *Management Support*. As shown in Table 50, this post hoc analysis indicated that neither input variable was significant at the α = 0.05 level when both were included in the model. However, *Management Support* had the smallest p-value (p = 0.07), indicating that it is a stronger – that is, more significant – predictor of *Tool Quality* than *Goal Clarity* is.

Table 50. Tool Quality on Goal Clarity and Management Support

Input Variable (X)	Coefficient (a)	Std. Error	p-value
Goal Clarity	.3389	.2371	.1529
Management Support	.4156	.2294	.0701

Table 51 summarizes the findings of the mediation analysis and the total strength of path for the mediation effects – i.e., a*b. Further implications of all these relations will be discussed in Chapter 5.

Table 51. Summary of Mediation Analysis Results for Overall Perceived Success

Input Variable (X)	Mediator Variable	Total mediated effect (a*b)	Partial or Full
Management Support	Tool Quality	.3161	Full

4.4.5 Mediation Analysis for % of Goals Met

There were no significant event process predictors for *% of Goals Met* in the continuous form. Therefore, no tests of mediation are performed. In the dichotomous variable analysis, there was one significant event process predictor: *Action Orientation*. Therefore, a mediation analysis was performed for those variables shown to have a significant relationship to *Action Orientation* (see Table 44). The regression of Goal Achievement – i.e., the dichotomous version of *% of Goals Met* – on *Action Orientation* and the predictor (X) was accomplished using logistic regression in GEE. As shown in Table 52, there was no indication that *Action Orientation* mediated the effects of these variables – i.e., there was no support for the mediation hypothesis for any of the input variables. In fact, for *Goal Difficulty*, the logistic regression did not produce any standard error estimates, indicating an error in the fit of the model to the underlying data – i.e., likely complete separation or lack of combinations (see Field, 2005).

Table 52. Goal Achievement (Dichotomous) on Internal Processes and Input Variable (X) Regressions (path b and path c')

Input Variable (X)	Coefficient (c')	Std. Error	p-value	Action Orientation Coefficient (b)	Action Orientation Std. Error	p-value
Goal Difficulty	-994.917	--	--	-669.366	--	--
Team Autonomy	.1014	.6085	.8676	-.6400	.4403	.1461
Work Area Routineness	-.6012	.3070	.0502	-.2921	.4615	.5267

4.5 Summary of Results of Hypothesis Tests

Table 53 summarizes the results of the tests of H1 – H10.

Table 53. Summary of Results of Tests of H5 – H10

Hypothesis	Findings	Overall Conclusion
H1: Social system outcome variables will be significantly correlated at the team level. H_0: Social system outcome variables are not significantly correlated at the team level.	• AT and TKSA, r = 0.710, p < 0.0001	Supported
H2: Social system outcomes will occur primarily at the team level, rather than individual level, indicated by significant intraclass correlation for social system outcome variables H_0: The intraclass correlation for social system outcomes is not significant	• AT, ICC(1) = 0.300, p < 0.0001 • TKSA, ICC(1) = 0.121, p < 0.0001	Supported
H3: Technical system outcome variables will be significantly correlated at the team level. H_0: Technical system outcome variables are not significantly correlated at the team level.	• IMA and OVER, n.s. • IMA and % of Goals Met, n.s. • OVER and % of Goals Met, n.s.	Not Supported
H5: Event input factors will be positively related to social system outcomes at the team level. H_0: Event input factors are not positively related to social system outcomes at the team level.	• For Attitude, Team Functional Heterogeneity and Management Support were significant direct predictors • For Task KSA, Goal Difficulty, Team Autonomy, Team Kaizen Experience, Team Leader Experience and Work Area Routineness were significant direct predictors	Partially Supported
H6: Event process factors will be positively related to social system outcomes at the team level. H_0: Event process factors are not positively related to social system outcomes at the team level	• For Attitude, Internal Processes was a significant direct predictor • For Task KSA, Internal Processes and Affective Commitment to Change were significant direct predictors	Partially Supported
H7: Event input factors will be positively related to technical system outcomes at the team level. H_0: Event input factors are not positively related to technical	• For Impact on Area, Team Autonomy and Management Support	Partially Supported

system outcomes at the team level.	• For Overall Perceived were significant direct predictors • Success no event input factors were significant direct predictors • For % of Goals Met, Team Kaizen Experience and Event Planning Process were significant direct predictors using the continuous response variable; in addition, Goal Difficulty, Team Leader Experience were significant direct predictors using the dichotomous response variable	
H8: Event process factors will be positively related to technical system outcomes at the team level. H_0: Event process factors are not positively related to technical system outcomes at the team level.	• For Impact on Area, Action Orientation was a significant direct predictor • For Overall Perceived Success, Tool Quality was a significant direct predictor • For % of Goals Met, no event process factors were significant direct predictors using the continuous response variable; however, Action Orientation was a significant direct predictor using the dichotomous response variable	Partially Supported
H9: Event process factors will partially mediate the relationship of event input factors and social system outcomes at the team level. H_0: Event process factors do not mediate the relationship of event input factors and social system outcomes at the team level.	• For Attitude, Internal Processes fully mediates Goal Clarity • For Task KSA, Internal Processes fully mediates Goal Clarity	Partially Supported
H10: Event process factors will partially mediate the relationship of event input factors and technical system outcomes at the team level. H_0: Event process factors do not mediate the relationship of event input factors and technical system outcomes at the team level.	• For Impact on Area, Action Orientation fully mediates Goal Difficulty and Work Area Routineness, and partially mediates Team Autonomy • For Overall Perceived Success, Tool Quality fully mediates Management Support	Partially Supported

4.6 Post-Hoc Control Variable Analyses

Following the development of the final regression models, post-hoc analyses were performed to determine whether any of the variation not accounted for in the final regression models could be accounted for by the inclusion of one or more “control” variables. These “control” variables were measured during data collection, but were not explicitly tested in the main analysis – i.e., as event input factors and event process factors – because they were not believed to be key variables influencing event outcomes (see Figure 1). However, it is likely that these “control” variables may be related to some of the event input and event process factors studied. The output of these post-hoc analyses are used to evaluate the robustness of the final regression models – i.e., stability under different levels of the “control” variables – and to evaluate whether any of the “control” variables appear promising for future research – i.e., as potential event input or event process predictor variables.

As is common in team research, one post-hoc analysis focused on determining whether *Team Size* had a significant direct effect in the final regression models. As mentioned, it was hypothesized that *Team Size* would not have a significant direct effect in the final regression models. A second post-hoc analysis was similarly used to determine whether *Event Duration* had a significant direct effect in the final regression models. It was hypothesized that *Event Duration* would not be one of the key variables influencing event outcomes, although it may be interrelated with some of the studied variables – e.g., team perceptions of *Goal Difficulty*, *Action Orientation*, etc. A third post-hoc analysis was used to determine whether the categorical variable *Event Type* – i.e., “implementation” versus “non-implementation” – had any significant effects on outcomes that were not accounted for by the final set of predictor variables in each model. It seems likely that “implementation” versus “non-implementation” events might have differed on some of the independent variables studied in this research – e.g., *Action Orientation*. However, the goal of this post hoc analysis was to determine whether there were any unmeasured variables related to *Event Type* that had significant relationships to the outcomes after controlling for the final set of predictors in each regression model. It should also be noted here that a related post-hoc analysis that was considered was examining the effects of a second categorical variable, *Event Methodology* – i.e., general process improvement, SMED, VSM, 5S, TPM. However, while there were 35 events in the general process improvement category, there were five or fewer events in each of the other four categories, and the VSM category only had two events. Therefore, there was insufficient sample size to complete a post-hoc analysis based on *Event Methodology*. However, this analysis is of interest in future research.

The fourth post-hoc analysis focused on testing the effect of *Team Kaizen Experience Heterogeneity* in the final regression models. *Team Kaizen Experience* (average team member experience with Kaizen events) was tested as a predictor variable in the current research; however, *Team Kaizen Experience Heterogeneity* was not. This analysis was used to evaluate the decision not to include *Team Kaizen Experience Heterogeneity* in the initial research model, as well as to evaluate whether studying *Team Kaizen Experience Heterogeneity* in future research would be of interest. The final (fifth and sixth) post-hoc analyses focused on the *Number of Main Goals* and the total *Number of Goals*, respectively. These analyses focused on determining whether these variables had significant relationships to outcomes after controlling for the final sets of predictors in the regression models. These analyses were of particular interest since the quantity of goals could (theoretically) be one measure of overall event scope and might potentially help explain some of the unexpected results for *% of Goals Met* – i.e., the negative relationship between *Team Kaizen Experience* and *Team Leader Experience* and *% of Goals Met*. The following paragraphs describe the results of the post-hoc analyses.

In the first post-hoc analysis, *Team Size* was not significant in any of the regression models using either the GEE "model based" standard error estimates or the OLS standard error estimates. For two of the models – *Impact on Area* and *% of Goals Met* in continuous form – the GEE "empirical" standard error estimates would suggest that *Team Size* had a significant direct effect. However, as mentioned earlier in this chapter, there is reason to believe that the GEE "empirical" standard error estimates are downwardly biased due to the small sample size at the organizational level. In both models, the reported effect of *Team Size* was negative and small ($\widehat{\beta}_{GEE}$ = -0.015 for both models). Overall, this post-hoc analysis suggests that *Team Size* does not have a significant direct effect on any of the outcome variables that is not accounted for by the variables already in the final regression models. In particular, *Team Size* does not appear to have a significant direct relationship to either of the two social system outcome variables.

In the second post-hoc analysis, *Event Duration* was not significant for any of the models using the GEE "model based" standard error estimates or the OLS standard error estimates. However, the p-value for the reported effect of *Event Duration* on *Task KSA* was fairly small ($p = 0.066$), using the GEE "model based" standard error estimates. Similarly, the GEE "empirical" standard error estimates would have been significant only for *Task KSA*. The sign of the reported effect of *Event Duration* on *Task KSA* was positive and relatively small ($\widehat{\beta}_{GEE}$ = 0.054).

Overall, the post-hoc analysis suggests that *Event Duration* does not have a strong, significant effect on any of the outcome variables that is not accounted for by the variables already in the final regression models. However, there is some weak support for the proposition that *Event Duration* has a unique, direct relationship to *Task KSA*. This relationship is logical. It is theoretically sound to suppose that longer events result in greater incremental gains in *Task KSA*, all else being equal, due increased length of exposure to problem-solving tools and concepts. This analysis suggests that perhaps *Event Duration* should continue to be analyzed in future research, particularly when *Task KSA* is the outcome variable of interest.

In the third post-hoc analysis, *Event Type* was not significant for any of the models using the GEE "model based" standard error estimates, the GEE "empirical" standard error estimates or the OLS standard error estimates. However, the p-value for the reported effect of *Event Type* on *Overall Perceived Success* was fairly small ($p = 0.079$), using the GEE "model based" standard error estimates. The sign of the reported effect of *Event Type* on *Task KSA* was positive and relatively large ($\widehat{\beta}_{GEE} = 0.532$), however with a large standard error. Overall, the post-hoc analysis suggests that *Event Type* does not have a significant effect on any of the outcome variables that is not accounted for by the variables already in the final regression models. However, there is some weaker support for the proposition that *Event Type* has a unique, direct relationship to *Task KSA*. The direction of the effect would suggest that implementation events are, in general, rated as more successful than non-implementation events. Therefore although the effect is not significant even for *Overall Perceived Success*, this analysis suggests that perhaps *Event Type* should continue to be analyzed in future research, particularly when *Overall Perceived Success* is the outcome variable of interest. This seems especially important since so little is known about the predictors of *Overall Perceived Success*.

In the fourth post-hoc analysis, *Team Kaizen Experience Heterogeneity* was not significant for any of the models using the GEE "model based" standard error estimates, the GEE "empirical" standard error estimates or the OLS standard error estimates. Overall, this post-hoc analysis suggests that *Team Kaizen Experience Heterogeneity* does not have a significant effect on any of the outcome variables that is not accounted for by the variables already in the final regression models. This supports the initial modeling decision not to consider *Team Kaizen Experience Heterogeneity* as a predictor variable – i.e., an event input factor.

In the final (fifth and sixth) post-hoc analyses, the GEE "model based" standard error estimates and OLS standard error estimates were not significant for either *Number of Main Goals* or *Number of Goals* in any of the regression models. However, the p-value for the reported effect of total *Number of Goals* on *Overall Perceived Success* was fairly small ($p = 0.064$), using the GEE "model based" standard error estimates. In addition, the GEE "empirical" standard error estimates would have been significant for the reported effect of total *Number of Goals* on both *Overall Perceived Success* ($p = 0.000$) and *% of Goals Met* in the continuous regression ($p = 0.030$). The sign of the reported effect of *Number of Goals* on *Overall Perceived Success* was positive and relatively small ($\widehat{\beta}_{GEE} = 0.088$), while the sign of the reported effect of *Number of Goals* on *% of Goals Met* was positive and small ($\widehat{\beta}_{GEE} = 0.015$). Overall, these post-hoc analyses suggest that *Number of Main Goals* and total *Number of Goals* do not have significant effects on any of the outcome variables that are not accounted for by the variables already in the final regression models. However, given the fairly small p-value for the effect of *Number of Goals* on *Overall Perceived Success* ($p < 0.10$), *Number of Goals* may be of continuing interest in future research on factors related to *Overall Perceived Success*. The direction of the effect indicates that events with a larger total *Number of Goals* had a higher *Overall Perceived Success*. Perhaps having a clear division between different objectives (i.e. more specific goals) helped the teams formulate better (i.e. more effective) problem-solving strategies. Or, perhaps a larger total *Number of Goals* meant that the facilitator had more explicitly defined facilitator (and/or management) expectations for the team, and thus had a more defined framework against which to measure *Overall Perceived Success*, as well as increased ability to help the team develop effective strategies (due to increased understanding of event objectives). Although *Goal Clarity* was measured in the current research, *Goal Clarity* measures the degree to which the team objectives were understandable, not the degree to which the team objectives comprehensively defined (all) stakeholder expectations.

Thus, none of the post-hoc analyses indicated clearly significant effects. However, some post-hoc analyses suggested avenues for future research. For instance, *Event Type* and *Number of Goals* might possibly be related to *Overall Perceived Success* and this potential relationship should be examined more extensively in future research.

CHAPTER 5: DISCUSSION

The following chapter provides more interpretation of the study results. First, the observed relationships between outcome variables are discussed. Next, the results of the regression modeling processes are discussed to refine hypotheses about the relationships between event input factors, event process factors and event outcomes. Both direct effects and indirect effects are discussed. Finally, the chapter concludes with discussion of the limitations of the present research.

One general note of caution is inserted here prior to discussing results as it applies to all the results discussed below. Due to the observational, cross-sectional nature of this study, the direction of causality is hypothesized, based on theory and the nature of the measures – e.g., the fact that the outcome measures specifically reference the impact of “this Kaizen event” – not conclusively proven empirically, which would require a controlled experiment. This is particularly true of the mediation models, which require a theoretical, hypothesized direction of causality for two sets of relationships. However, the argument for the hypothesized direction of causality is noticeably strengthened for variables measured through the Kickoff Survey (*Goal Clarity*, *Goal Difficulty*, and *Affective Commitment to Change*), since this measurement was taken before the Report Out Survey, and for the objective variables measured through the surveys or the Event Information Sheet (*Team Functional Heterogeneity*, *Team Kaizen Experience*, *Team Leader Experience*, and *Event Planning Process*). In the case of these variables, particularly the objective measures, it is difficult to argue for reverse causality – i.e., that the outcomes in fact caused the measured values of these “independent” variables – unless the outcomes are assumed to be pre-existent to the event and global – i.e., known and shared by the facilitator or other organizational employee planning the event. While this argument cannot be entirely ruled out due to the non-experimental nature of the study, both the timing of the measurements and the nature of the outcome measures – i.e., reference to the specific Kaizen event – make this proposition highly unlikely. Although the argument that the perceptual variables measured through the Report Out Survey and the Event Information Survey are also antecedent to outcomes is relatively strong theoretically, there is more potential for contamination of causality effects for these measures, due to the fact that they are perceptual measures that were measured concurrently with the outcomes. Finally, it should also be noted that, in observational studies, there is also always some risk that the statistical relationship exists because both outcomes and predictors are correlated with the true, unknown causes of the outcomes, but that the predictor variable does not in any way determine the level of that cause.

5.1 Relationship between Kaizen Event Outcomes

Although not specified as testable hypotheses, two of the questions of interest in this research focus on measuring the outcomes of Kaizen events (see the first two research questions in Section 1.2 and the first two research purposes in Section 1.3), since the range of Kaizen event outcomes within and across organizations is rarely reported in the Kaizen event practitioner literature. Part of this objective was accomplished by defining and testing appropriate measures of event outcomes, particularly social system outcomes – i.e., *Attitude* and *Task KSA*. A related investigation is the examination of the range of response for each dependent variable in the study. Finally, these research questions can also be addressed through the investigation of the interrelationships between social system outcomes and the interrelationships between technical system outcomes.

The events studied in this research cannot be considered a true random sample. Although the events studied within each organization were randomly selected, the organizations were not randomly selected. Thus, the results observed here cannot be taken as estimates of the expected results from all organizations conducting Kaizen events – i.e., true population estimates. In addition, the boundary criteria employed to select organizations were intended to identify organizations that were fairly mature in their Kaizen event program and had been conducting Kaizen events for some time. However, although not a population estimate, the range of outcomes identified in this research lends support to the hypothesis that even organizations that are mature in their Kaizen event programs experience some variation in outcomes. The overall outcomes from the participating organizations are presented in Appendix S.

For the participating organizations, Kaizen event teams reported consistently positive perceptions of human resource (social system) outcomes from Kaizen events. The mean team response was 5.00 (“agree”) for *Attitude* and 4.87 for *Task KSA*. In addition, the minimum observed response was close to 4.0 (“tend to agree”) for both *Attitude* and *Task KSA*. The overall range of response for *Attitude* was 4.00 – 5.83, while the overall range of response for *Task KSA* was 3.94 – 5.63. These results suggest that, for the organizations in the study, Kaizen events were associated with positive employee perceptions of the impact of the events on their *Attitude* toward (liking for) Kaizen events and their *Task KSA*. However, it should also be noted here that, although the Kaizen events studied in this research were randomly selected for inclusion, organizations appear to use a non-random process in selecting team members for Kaizen events – based on employee personality, technical expertise, etc. Thus, there may be some unmeasured factor(s) that may have caused the team members to be more likely to report positive outcomes than other employees in the organizations, and therefore there is a possibility that the results may not generalize to a team composed of randomly selected employees. As indicated in Appendix S, there was no significant difference across participating organizations in terms of *Attitude* (p = 0.305), although there was a significant difference in terms of *Task KSA* (p = 0.009). (Note, this effect is not significant if a Bonferroni correction is employed).

Similarly, for the participating organizations, Kaizen event teams reported positive perceptions of the impact of their activities on the target work system (technical system impact). The mean team response for *Impact on Area* was 4.91. In addition, the minimum observed response was slightly less than 3.5, which is the midpoint of the survey scale (a neutral response range). The overall range of response for *Impact on Area* was 3.48 – 5.78. This range suggests that, for the organizations in the study, Kaizen events were generally associated with positive employee perceptions of the impact of the events on the target system. However, as described previously, due to the non-random team selection processes within the participating organizations, it is possible that some unmeasured characteristic made these employees more likely to report positive results that other employees in the organizations. As indicated in Appendix S, there was no significant difference across participating organizations in terms of *Impact on Area* (p = 0.436).

In addition, for the participating organizations, Kaizen event facilitators reported positive perceptions of overall event success (*Overall Perceived Success*). However, there was a wider range of variation in response. The median response was 5.0 ("agree"); however, the minimum response was 1.0 ("strongly disagree"), indicating that the data set did contain some events that were viewed as substantially less successful from the facilitator's perspective. In total, three out of the 51 events studied were viewed as less successful, receiving a rating of 1.0 or 2.0 on the six-point response scale. The overall range of responses was 1.0 - 6.0. These results indicate that, for the participating organizations, Kaizen event facilitators viewed most, although not all, of the events in the data set as at least somewhat successful. As indicated in Appendix S, there was no significant difference across participating organizations in terms of *Overall Perceived Success* (p = 0.588).

% of Goals Met was more variable than the other four outcomes measures. Most of the Kaizen event teams in the sample (35) were successful in meeting 100% of their main goals. However, for the remaining minority of events, success was more variable. Eight teams (approximately 16% of the sample) achieved 50% or less of their goals. In addition to the eight teams that achieved 50% or less of their goals, an additional five teams achieved

between 50-85% of their goals. Thus, while the majority of the teams were completely successful in meeting their specified goals (at least in terms of short-term results), a substantial minority fell noticeably short of their specified goals. As mentioned, these results are quite different than most of the published, anecdotal accounts, where primarily successful teams are highlighted. These results indicate that, for the participating organizations, which were relatively experienced in using Kaizen events, success was still variable in terms of goals achievement. Slightly more than 25% of the events studied fell noticeably short of completely achieving their specified goals -- i.e., had a goal achievement of 85% of less. Due to the nature of the study, these results cannot be directly extrapolated to other organizations, but they do support the hypothesis that some variability in goal achievement exists even within organizations that are relatively more mature in their Kaizen event programs. Additional research will be needed to discover whether similar results exist for other organizations.

Observing the levels of outcome variables, while related to the objectives of this research, does not represent a set of testable hypotheses. Instead, the testable hypotheses specified in this research (H1, H2, H3 and H4) focused on the interrelationships between outcome variables. As hypothesized (see H1), and suggested in the Kaizen event practitioner literature (see Chapter 2), employee affect toward participation in Kaizen events (*Attitude*) and employee belief that the Kaizen event strengthened their continuous improvement knowledge, skills and work motivation (*Task KSA*) were highly related (see Table 19). However, the relationship was not perfect – only about 50% of the variance is shared (see Table 19). This supports the hypothesis that the two variables should be managed separately by organizations seeking to create these favorable human resource outcomes. Further supporting this point, *Attitude* and *Task KSA* were found to have a unique set of predictors (see Table 33 and Figures 2 and 3). There were some common predictors of both outcomes – namely *Internal Processes* (direct for both) and *Goal Clarity* (indirect for both). However, *Attitude* had two unique predictors not shared with *Task KSA*: *Team Functional Heterogeneity* (direct) and *Management Support* (direct). In addition, *Goal Difficulty*, *Team Autonomy*, *Team Kaizen Experience*, *Team Leader Experience*, *Work Area Routineness*, and *Affective Commitment to Change* were significant (direct) predictors for *Task KSA* but not for *Attitude*. Finally, both variables were found to occur at the group level, rather than individual level, supporting the hypothesis that learning in Kaizen events occurs at the group level (H2), which is consistent with learning theory.

Contrary to the study hypotheses (H3), there were no significant relationships between the different measures of technical success – *Impact on Area* (an event impact rating provided by the team), *Overall Perceived Success* (rated by the event facilitator), and actual *% of Goals Met*. This result is somewhat unexpected, although not perhaps as counterintuitive as it initially appears, due to the (intended) differences in the focus of the measures. *Impact on Area* focused on the extent to which employees believed the event had resulted in improvement in the target system -- i.e., work area. Since this measure does not directly address goal achievement, it is possible that an event which failed to achieve its goals – e.g., if the goals were very ambitious – could still have resulted in clear and measurable improvement in the target system, resulting in a relatively high team score for *Impact on Area*. However, the same event evaluated on the *% of Goals Met* measure could have looked much less favorable in terms of technical success – perhaps even as a complete failure). Meanwhile, a team could achieve its goals without resulting in (much) visible improvement in the target work area, potentially resulting in the reverse situation.

Overall Perceived Success, rated by the facilitator, reflects the facilitator's holistic expert judgment about the overall effectiveness of the event.[12] For given event, this could be based on a weighting of several factors, which could be different across events and across facilitators. From qualitative comments provided in the Event Information Sheet, it appears that one influential factor is likely facilitator perceptions of stakeholder satisfaction – i.e., top management approval of the team's solution. This was illustrated by the case of one event in the data set where *% of Goals Met* and *Impact on Area* were both high and the facilitator rating of *Overall Perceived Success* was extremely low (1 = "strongly disagree") (see Farris et al., 2006 for more details). The facilitator stated in the Event Information Sheet comments that this was due to the fact that top management had rejected the team's solutions. However, additional research would be necessary to determine whether this relationship exists across other events and organizations.

It should also be noted that the distributional properties of *% of Goals Met* and *Overall Perceived Success* may also have contributed to failure to find significant correlations between the outcome variables. *% of Goals Met* was highly skewed, even following a logarithmic transformation. This was due to the fact that so many of the teams (35 out of 51) achieved all of their improvement goals and thus had the same score for *% of Goals Met*. Meanwhile, *Overall Perceived Success* was symmetric, but highly truncated.

Overall, the lack of strong correlation between technical system outcome measures supports the proposition that a holistic set of technical success measures are needed to fully assess the technical success of a given event, both for

[12] The facilitator, in most cases, is an organizational employee who spends a large proportion of his or her time planning and conducting Kaizen events; generally, for the organizations studied, the facilitator had conducted several previous events, providing a broad knowledge basis for comparison.

researchers and managers. An event may look successful in terms of *% of Goals Met*, but might miss receiving buy-in from critical stakeholders or fail to have an immediate, noticeable positive impact on the target system. In addition, team member perceptions of event impact do not always coincide with facilitator perceptions of event success. This could result in a organizational disconnect that must be addressed to prevent Kaizen event team members from becoming disillusioned with the Kaizen event process if management does not ultimately (fully) accept and support their solutions (Lawler & Mohrman, 1985, 1987).

Despite the lack of direct relationship, the technical system outcomes did share some common predictors. *Management Support* was a common predictor of both *Impact on Area* (direct) and *Overall Perceived Success* (indirect). Meanwhile, *Goal Difficulty* and *Action Orientation* were common predictors of both *Impact on Area* and *% of Goals Met*. *Goal Difficulty* was a direct predictor of *% of Goals Met* and an indirect predictor of *Impact on Area* (through *Action Orientation*). *Action Orientation* was a direct predictor of both measures; however, the sign of the relationship was different for the two measures. *Action Orientation* had a positive relationship to *Impact on Area* and a negative relationship to *% of Goals Met*. As expected, events with a higher *Action Orientation* – i.e., relatively more time spent in the target work area versus "offline" in meeting rooms – also reported higher team member perceptions of event *Impact on Area*. However, increased *Action Orientation* was not associated with increased goal achievement (*% of Goals Met*). Instead, *Action Orientation* had a significant negative relationship to *% of Goals Met* (in the dichotomous form), suggesting that too much focus on "hands on" activities may inhibit goal achievement. Although these results appear counterintuitive at first, this set of relationships is possible due to the lack of correlation between the two technical system outcomes.

Finally, there was some support for a relationship between technical system outcomes and social system outcomes (H4). *Impact on Area* was highly correlated with both *Attitude* and *Task KSA*. Approximately 40% of the variance is shared between *Impact on Area* and *Attitude*, and approximately 47% of the variance is shared between *Task KSA* and *Attitude*. This lends support to the hypotheses in the Kaizen event practitioner literature, as well as the organizational change literature, that immediately visible results (*Impact on Area*) are associated with increased team member motivation to participate in events (*Attitude*). This relationship is hypothesized to be due, at least in part, to the fact that participating employees can see immediate return on their investment – i.e., evidence that the methodology works in producing improvements. However, the direction of causality between *Attitude* and *Impact on Area* is hypothesized only based on related literature and cannot be validated due to the cross-sectional nature of this study. It is possible that a follow-up longitudinal study with multiple measurements on *Attitude* and *Impact on Area* during the event might be used to more directly study this relationship. However, testing effects due to multiple administrations of the same survey instrument within a very short time window (less than one week) must be considered in any proposed study design. The relationship between *Task KSA* and *Impact on Area* lends support to the proposition in Kaizen event practitioner literature that Kaizen events are an effective, “hands-on” training tool for equipping participating employees with knowledge, skills and motivation related to continuous improvement, by allowing them to immediately apply those new KSAs to the problem at hand. One practitioner article even refers to Kaizen events a type of “just-in-time” training (Drickhamer, 2004a). It seems likely that the relationship between *Task KSA* could be at least somewhat reciprocal. Increased team gains in *Task KSA* may be expected to contributed to increased *Impact on Area*, all else being equal, because teams with greater gains in *Task KSA* might be better able to apply the continuous improvement concepts. Similarly, events where teams are able to achieve a high *Impact on Area* may result in greater gains in *Task KSA*, because the meaning and purpose of problem-solving tools and concepts may be better understood when visibly and effectively applied. Again, the directionality of this relationship is hypothesized and represents a proposition for future research. It cannot be verified in this cross-sectional study. It is also interesting to note that one potential mediating variable – *Tool Quality* – showed no direct relationship to either *Impact on Area* or *Task KSA*.

On additional note in discussion of the relationship between technical and social system outcomes is the overlap in sets of predictors. *Management Support* was a significant predictor of both one social system outcome (*Attitude*) and two of the technical system outcome (*Impact on Area* and *Overall Perceived Success*). *Goal Difficulty* was also a significant predictor of one social system outcome (*Task KSA*) and two technical system outcomes (*Impact on Area*, *% of Goals Met*). However, the sign of the relationship was different for the two types of outcomes. *Goal Difficulty* had a positive relationship to *Task KSA* and a negative relationship to the two technical system outcomes (*Impact on Area*, *% of Goals Met*). The increased challenge from more difficult goals appeared to be beneficial in terms of team member gains in *Task KSA* but detrimental in terms of *Impact on Area* and team goal achievement (*% of Goals Met*). *Work Area Routineness* was also a significant (positive) predictor of one social system outcome (*Task KSA*) and one technical system outcome (*Impact on Area*). Similarly, *Team Autonomy* was a significant predictor of one social system outcome (*Task KSA*) and one technical system outcome (*Impact on Area*). Meanwhile, *Team Kaizen Experience* and *Team Leader Experience* were common (negative) predictors of one social system outcome (*Task KSA*) and one technical system outcome (*% of Goals Met*). The following sections discussed the relationships observed for each outcome variable in more detail.

5.2 *Significant Predictors of Attitude*

Figure 2 depicts the overall model of significant relationships between *Attitude* and the predictor variables that were identified in this research. Table 54 contains the relative effect sizes – i.e., GEE regression coefficients – for the variables in the final model. Two words of caution should be inserted here, which also apply to the effect size tables in the following sections. First, effect sizes must be interpreted with caution due to the differences in scale of measurement across variables – i.e., a six-point interval scale for *Goal Clarity*, *Internal Processes* and *Management Support*, and an index from 0-1 for *Team Functional Heterogeneity*. The GEE regression coefficients are unstandardized. Second, the indirect effect especially must be interpreted with caution because it represents a raw (unmoderated) estimate of the mediated effect. Barring suppressor effects, it is likely that this represents an upper bound on the actual mediated effect However, the effect size table can still be useful for gauging the relative effect size of the different variables, particularly the variables measured on the same rating scale – i.e., determining which survey variables have the largest effect sizes.

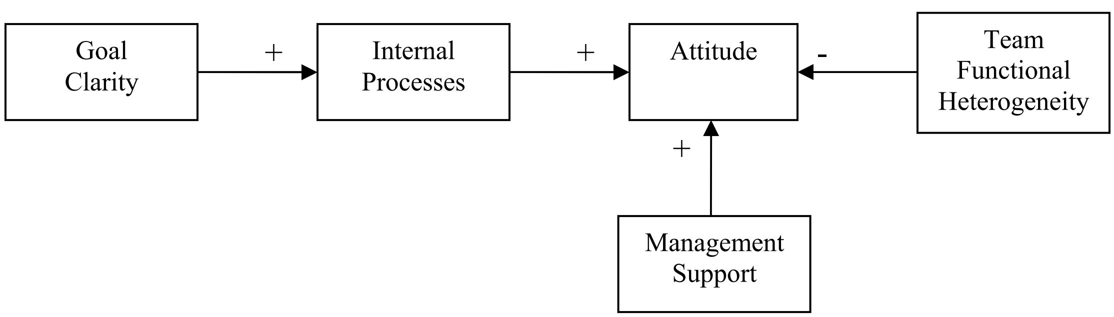

Figure 2. Overall Model for Significant Predictors of Attitude

Table 54. Effect Size Table for Attitude

Predictor	Direct Effect	Indirect Effect
Team Functional Heterogeneity	-.547	
Internal Processes	.694	
Management Support	.250	
Goal Clarity		.414

This research found that the most significant predictors of *Attitude* toward Kaizen events were *Team Functional Heterogeneity* (direct negative), *Management Support* (direct positive), *Internal Processes* (direct positive), and

Goal Clarity (indirect positive). In addition, as described in Chapter 4, there is some weaker support that *Management Support* and *Team Autonomy* may have positive indirect effects. However, due to the weaker support, these potential indirect effects are not discussed here.

Most of the variables demonstrated a positive relationship with *Attitude*; however, *Team Functional Heterogeneity* demonstrated a negative relationship with *Attitude*. More diverse teams were associated with lower levels of liking for Kaizen activities. This finding is consistent with previous team research suggesting that cross-functional teams may make better decisions (McGrath, 1984; Jackson, 1992; Jehn et al., 1999; Lovelace et al., 2001), but also experience lower levels of enjoyment in working together (Baron, 1990; Ancona & Caldwell, 1992; Amason & Schweiger, 1994). However, this finding is particularly interesting since the use of a cross-functional team is one of the most common recommendations in the Kaizen event literature (e.g., LeBlanc, 1999; Drickhamer, 2004b; Rusiniak, 1996; Demers, 2002; Smith, 2003; Cuscela, 1998; McNichols et al., 1999; Martin, 2004; Sheridan, 1997b; Vasilash, 1993; Vasilash, 1993; Adams et al., 1997; Melnyk et al., 1998; Sheridan, 2000a; Pritchard, 2002; Laraia, 1998; Harvey, 2004; Foreman & Vargas, 1999).

Perhaps surprisingly, *Team Functional Heterogeneity* did not show a significant relationship to team ratings of *Internal Processes* (tested during the mediation analysis). Assuming the direction of causality in the research model is correct, *Team Functional Heterogeneity* appears to have acted only directly on *Attitude* toward Kaizen events versus directly and indirectly through *Internal Processes*. It may be that, in more homogeneous teams, team members received more direct enjoyment from working with their peers – i.e., people they work closely with on a day-to-day basis – than they did in working in more cross-functionally diverse teams, where they were less likely to know other team members as well. Rather than enjoyment from working with those who are similar, this finding could instead reflect a lower affect for working with people who are different or less well known – i.e., from different cross-functional backgrounds. This second proposition is similar to the first proposition above; however, the first proposition focuses on enjoyment gained from working with "friends" – i.e., people who are well known to each other – whereas the second proposition focuses on potential discomfort in working with people coming from different backgrounds and schools of thought. Either effect would not necessarily be reflected in *Internal Processes* scores, since *Internal Processes* measures whether there was any evidence of disharmony (lack of respect and breakdowns in communication) within the group, rather than the extent to which team members enjoyed working together. This also agrees with previous research which has found that diverse teams may develop very effective communication mechanisms (Azzi, 1993; Earley & Mosakowi, 2000). Support for either proposition could be analyzed by including additional measures in future research – i.e., a direct measure of "team spirit," a scale measuring comfort in working with those well known to each other, a scale measuring comfort in working with those from different backgrounds.

The findings of this research with regard to the apparent negative impact of *Team Functional Heterogeneity* on *Attitude* are not at this point interpreted to imply that organizational managers should stop using cross-functional teams in Kaizen events, since other research (as well as team member and facilitator written comments from the Kaizen events studied in this research) suggests that using a cross-functional team can improve solution quality. For problems that cut across different functional boundaries, cross-functional teams are still likely particularly appropriate, in part to help create open channels of communication across the different functions. Also, as reported in Appendix S and Section 5.1, despite the negative relationship between *Team Functional Heterogeneity* and *Attitude*, all of the teams studied in this research reported at least somewhat positive *Attitude* outcomes. What these findings may imply, if they hold across additional organizations, is that managers should recognize that increasing team diversity might lower team member enjoyment of Kaizen events, which ultimately *could* lower employee buy-in and enthusiasm for Kaizen events, if counter measures are not taken. Thus, in events where a more functionally diverse team is used, the facilitator may want to pay particular attention to maintaining positive *Internal Processes*, *Goal Clarity*, *Team Autonomy* and *Management Support*, in order to counteract the potentially negative impact of increased functional diversity, provided the hypothesized direction of causality is correct. These countermeasures would seem particularly important for teams with low average *Team Kaizen Experience*, since it would seem particularly important for creating buy-in to the Kaizen event program for employees to enjoy their first experiences with Kaizen events.

Internal Processes was another significant, direct predictor of *Attitude* toward Kaizen events. This result has strong face validity, since *Internal Processes* measures the harmony of the team with regard to respect for persons and open communication. Since the *Attitude* scale references the current Kaizen event, it seems logically difficult to imagine a situation where team members would report positive *Attitude* toward Kaizen events, while reporting negative *Internal Processes*. Conceptually, positive *Internal Processes* seem to be a necessary (but not sufficient) condition for positive *Attitude* toward events. *Internal Processes* is a high leverage variable in that this research also suggests it can be influenced in several ways. Three additional variables that this research suggests can be used to improve team *Internal Processes* will be discussed presently. It is also noted here that the event facilitator would appear to play a crucial role in directly enabling positive *Internal Processes*. The facilitator often has the primary organizational role in selecting the personnel for events (i.e., a good set of "team players"), establishing ground rules (such as respect for persons, open communication, etc.) and keeping team discussions "on track" during the event. This is suggested both by team member written comments from the events studied and by the Kaizen event practitioner literature (e.g., Mika, 2002).

In addition to his/her direct role in the facilitation process, this research suggests that there is at least one other variable that the facilitator or other organization personnel can manipulate to impact *Internal Processes* and, thereby, *Attitude* toward Kaizen events, provided the hypothesized direct of causality is correct. The mediation analysis results suggest that *Goal Clarity* is a significant, positive indirect predictor of *Attitude* through *Internal Processes*. *Goal Clarity* and *Internal Processes* were measured at two distinct points in time – at the Kickoff meeting and the Report Out meeting, respectively. Thus, there is reduced likelihood of confounding between the two variables and there is greater support that the hypothesized direction of causality – i.e., *Goal Clarity* causing *Internal Processes* – is correct. *Goal Clarity* measures the initial team member perceptions of the extent to which expectations for the event have been well defined, such that they are understood by all team members at the time of the Kickoff meeting. It appears that *Goal Clarity* could positively impact *Internal Processes* in at least two related ways. First, by clearly defining team goals prior to the start of the event, valuable event time which must otherwise be spent refining the goals (Letens et al.., 2006) is "freed-up." Although recommended in some Kaizen event practitioner resources, Letens, Farris and Van Aken (2006) suggests that attempting to define (or refine) team goals during the event can be a stressful and frustrating (i.e., non-harmonious) exercise. Second, clearly defining goals provides focus and a common language of communication for the team – thus likely facilitating open communication among team members, as well as enabling team members to more fully consider other team member's ideas since they would be better able to see how the ideas relate to event objectives. These findings support the hypothesis that facilitators and others in organizations who plan events should not shortcut the process of clearly defining (and refining) team objectives prior to the event. These findings also agree with previous research which has found significant relationships between *Goal Clarity* and team effectiveness (e.g., Van Aken & Kleiner, 1997; Doolen et al, 2003a). In addition to the standard event charters provided in published Kaizen event guidebooks (facilitator manuals), there are several specific recommendations for developing clear goals in the

Kaizen event practitioner literature (see Table 1). Also, there are several innovative practices in use in organizations, including the use of pre-event "sensing sessions" (i.e., focus groups) where stakeholder groups provide feedback used to refine the goals in a pre-event working session (Farris & Van Aken, 2005). Finally, this finding suggests that some of the suggestions in the Kaizen event practitioner literature, such as having the team members develop the goals (e.g., Wittenberg, 1994*)* and even determine the target work area during the event (e.g., Kumar & Harms, 2004) may not be advisable.

Finally, *Management Support* had a positive direct relationship to *Attitude*. Thus it appears that *Management Support* appears to either impact *Attitude* directly or through some unmeasured process factor. The final set of questions in the *Management Support* construct relate to adequacy of materials, equipment and supplies and help from others in the organization. It seems likely that increased resource support – i.e., the current *Management Support* scale – is one method of communicating overall management support, as well as management concern for the well-being of the team, to the Kaizen event team members. Increased perceptions of being involved in work that is important to the organization – i.e., events that have high overall management support – could result in more positive *Attitude* toward Kaizen events. Again, this link is hypothesized and would need to be tested in future research. For instance, in future research direct measures of perceived event importance to management could be used – i.e., event priority, as well as event scope/size. It could also be that increased levels of positive *Attitude*s are due not primarily to increased perceptions of event importance to management but, rather, as mentioned above, to increased employee perceptions that management is concerned for their individual well-being – i.e., the feeling of being valued by management. To test this proposition, a scale measuring team member perceptions of management's concern for their well-being and/or respect for their contributions could be included in future research. Or, it could be, more directly, that team members tend to enjoy events where they do not experience resource shortfalls more than events where there are resource problems. Given the importance of *Management Support* – the fact that it has a positive relationship to three out of five outcome variables – additional research on the exact nature of the relationship between *Management Support* and *Attitude* is needed. It should also be noted that none of the three propositions discussed above are mutually exclusive – i.e., all could have some simultaneous impact on *Attitude* – although some order of relative importance would be likely to emerge in future research.

5.3 *Significant Predictors of Task KSA*

Figure 3 depicts the overall model of significant relationships between *Task KSA* and the predictor variables, which were identified in this research. Table 55 contains the relative effect sizes – i.e., GEE regression coefficients – for the variables in the final model.

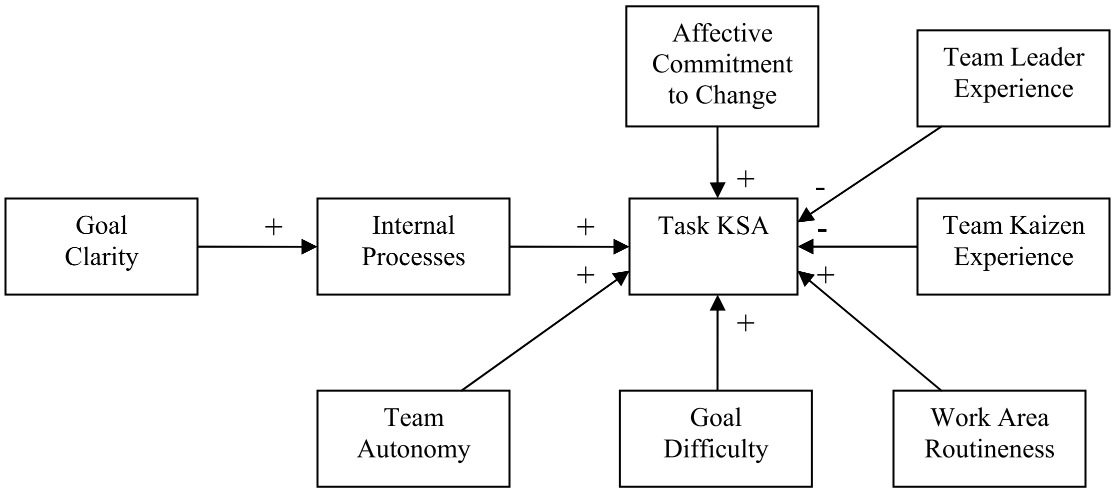

Figure 3. Overall Model for Significant Predictors of Task KSA

Table 55. Effect Size Table for Task KSA

Predictor	Direct Effect	Indirect Effect
Goal Difficulty	.119	
Team Kaizen Experience	-.398	
Team Leader Experience	-.195	
Work Area Routineness	.094	
Affective Commitment to Change	.222	
Internal Processes	.465	
Team Autonomy	.234	
Goal Clarity		.316

Of the outcome variables studied in this research, by far the variable with the largest number of significant predictors was *Task KSA*, suggesting a large number of ways that *Task KSA* could potentially be positively impacted, but also increased complexity of management of this outcome variable, providing the results of this research hold across additional organizations. This research found that the most significant predictors of team member gains in *Task KSA* were *Goal Difficulty* (direct positive), *Team Kaizen Experience* (direct negative), *Team Leader Experience* (direct negative), *Work Area Routineness* (direct positive), *Team Autonomy* (direct positive), *Internal Processes* (direct positive), *Affective Commitment to Change* (direct positive), and *Goal Clarity* (indirect positive). In addition, as described in Chapter 4, there is some weaker support that *Management Support* and *Team*

Autonomy may have positive indirect effects. However, due to the weaker support, these potential indirect effects are not discussed here. Again, the *Task KSA* variable measures the perceived extent of team members' incremental gains in continuous improvement KSAs as a result of the current event.

Goal Difficulty had a positive relationship to *Task KSA*. This result has strong face validity. Group and individual learning theory suggests that, in general, more complex – i.e., less routine – problems are likely to be associated with increased learning, due to increased stimulation of individual and team creativity processes (e.g., Amabile, 1989; Druskat & Pescosolido, 2000; George & Zhou, 2001). In addition, increased goal difficulty has been found to result in greater effort and performance (Locke & Latham, 1990; Locke & Latham, 2002), as well as greater learning (Wood & Locke, 1990). However, although this relationship is rarely qualified, it does seem likely that it may only hold up to a certain point; if a problem becomes too complex, it would seem likely that very little specific learning would occur due to team inability to effectively respond to the situation (Atkinson, 1958; Erez & Zidon, 1984). Thus, the finding that, in the current research, more difficult goals are associated with greater team member gains in *Task KSA* appears to agree with previous research which associates greater learning with more challenging problems. From a social system (human resource development and training) perspective, these findings agree with recommendations in the Kaizen event practitioner literature that Kaizen event goals should be challenging (e.g.. LeBlanc, 1999; Minton, 1998; Rusiniak, 1996; Cuscela, 1998; Bradley & Willett, 2004; Bicheno, 2001; Tanner & Roncarti, 1994; Treece, 1993; Kumar & Harms, 2004; Gregory, 2003). However, for the events studied in this research, there appears to be a tradeoff between gains in *Task KSA* due to increased *Goal Difficulty* and decreased technical system results – specifically, lower *Impact on Area* and *% of Goals Met*. In the current research, *Goal Difficulty* had a positive relationship to *Task KSA* but a negative relationship to *Impact on Area* and *% of Goals Met*, as will be further discussed in the sections on *Impact on Area* and *% of Goals Met*.

For the events studied in this research, *Team Kaizen Experience* had a negative relationship to *Task KSA*. This result is perhaps counterintuitive until the nature of the *Task KSA* variable is considered. As mentioned, *Task KSA* actually measures team member perceptions of the *incremental* gains in *Task KSAs* among their team members as a result of the event, rather than the actual amount of *Task KSAs* possessed. Thus, *Task KSA* is an incremental versus an absolute measure. As research on learning has shown, the rate of increase in KSAs is greatest among the first few exposures to a problem, methodology, etc. – i.e., the classic "learning curve" (Wright, 1936). Thus, this finding that, all else being equal, the incremental gains in *Task KSAs* are lower for more experienced teams versus less experienced teams agrees with related theory. It is not clear that the relationship of *Team Kaizen Experience* and training outcomes – i.e., gains in *Task KSA* – has been explicitly considered in the Kaizen event practitioner literature. While it has been suggested that Kaizen events can be used as a training tool (e.g., Drickhamer, 2004a), if the findings of this research hold across additional organizations, the average Kaizen event experience of team members should be explicitly considered when Kaizen events are intended to be implemented as a training tool. These findings suggest that a Kaizen event with a team composed of all or mostly inexperienced employees may be more effective as a training tool than a Kaizen event team composed of several more experienced team members, all else being equal. However, there are many other variables (*Goal Difficulty*, *Work Area Routineness*, *Team Autonomy*, *Internal Processes*, *Affective Commitment to Change*, and *Goal Clarity*) that facilitators can manipulate to potentially counteract negative effects on overall team learning due to having a more experienced team. For instance, providing the team with more challenging goals (i.e. increased *Goal Difficulty*) and increased *Team Autonomy* could help compensate for increased average experience. Or, more positive *Internal Processes* could also act as a countermeasure to help stimulate learning in the presence of high *Team Kaizen Experience*. Although this research did not find any significant positive relationships between *Team Kaizen Experience* and any of the outcome variables, this does not preclude *Team Kaizen Experience* from having some unmeasured beneficial effects on these or other outcomes (i.e., direct effects, indirect effects or interaction effects). Based on the current state of knowledge after this research, it does not appear advisable to suggest that organizations try to reduce *Team Kaizen Experience* on all events. There would appear to be situations where a high *Team Kaizen Experience* may be warranted. For instance, perhaps in a situation where training cannot be provided to the team as part of the event or a situation where the most appropriate team members for the event – in terms of securing buy-in, process knowledge, positive attitudes, problem-solving skills, subject matter expertise, etc. – also happen to have a high degree of *Team Kaizen Experience*. However, it is recommended that countermeasures be taken to compensate for the apparent negative effects of *Team Kaizen Experience* on *Task KSA* and *% of Goals Met*.

It is an unexpected finding that, for the events studied in this research, *Team Leader Experience* also had a negative relationship to *Task KSA*. This finding would seem to warrant future research using quantitative survey tools and/or qualitative interviews to understand the exact nature of this relationship. One working hypothesis is that, in teams with a more experienced team leader, the solution and the solution development methods may be more driven by the team leader versus the team as a whole, resulting in lower average *Task KSA* for the team as a whole.

(Note, the *Task KSA* questions used the team as a whole, rather than the individual, as the referent). The situation of having an experienced team leader may therefore be analogous to having an expert in other decision making contexts. The team as a whole may defer to the team leader's expertise and therefore have lower average gains in *Task KSAs* versus the situation where there is not a strong expert and all team members would appear to share more of the burden of applying tools and concepts and thus learning how to apply them. As mentioned, this proposition has been developed based on the findings of this research and will require further investigation through future research.

Work Area Routineness had a relatively small, but significant, positive impact on *Task KSA*. Work areas that were more routine (i.e., less complex) were associated with greater employee gains in *Task KSA*. At first, this may seem to contradict the finding that *Goal Difficulty* (i.e., having more complex goals) was associated with increased *Task KSA*. However, complexity of the target problem (*Goal Difficulty*) and complexity of the target work area (*Work Area Routineness*) are distinct constructs. For instance, a target work area might be quite complex (e.g., an engineering process), but the goals set for the team might be relatively achievable (e.g., "create a current state value stream map"). Conversely, a target work area might be quite routine (e.g., a single, dedicated machine) but the goals set for the team might be challenging, stretch goals (e.g., "achieve a 90% reduction in setup time") However, a given goal (e.g. "achieve a 50% reduction in cycle time") would be expected to result in different levels of perceived *Goal Difficulty* when applied to work areas of different levels of Routineness – i.e., more complex work areas would be expected to have higher *Goal Difficulty* ratings. The positive relationship between *Work Area Routineness* and *Task KSA* is hypothesized to be due to the fact that the application of lean tools and concepts is more tractable and salient for more routine –i.e., repetitive and predictable – work areas (James-Moore & Gibbons, 1997; Cooney, 2002: Papadopoulou & Ozbayrak, 2005), while it is not always clear how lean concepts and tools can be applied to more complex work areas. Thus, it seems likely that team members reported higher *Task KSA* for more routine work areas, because they were able to more easily understand (i.e., learn) how lean concepts could be applied to the work area, whereas lean concept application may not have been as clear in more complex work areas. As Table 15 shows, several of the questions in the *Task KSA* construct focus on understanding of the philosophy of continuous improvement and how it can be applied to the target work area. Although this research found a positive relationship between *Work Area Routineness* and *Task KSA*, it does not appear that organizations should stop holding Kaizen events in more complex work areas. First, the effect is relatively small (i.e., $\widehat{\beta}_{GEE} < 0.10$). Second, more complex work areas, such as engineering processes, may represent the bottlenecks in overall organizational improvement (Murman et al, 2002). However, this finding suggests that, when events are being held in more complex work areas, facilitators should recognize the potential detriment to *Task KSA* and may want to employ countermeasures – e.g., increasing *Team Autonomy*, *Internal Processes*, *Affective Commitment to Change*, *Goal Clarity* and *Management Support*.

For the events studied in this research, higher *Team Autonomy* was associated with increased *Task KSA*. While the recommendations regarding *Team Autonomy* in the Kaizen event practitioner literature often relate to its supposed relationship to solution quality, these findings indicate that *Team Autonomy* appears to also be important from a human resource development perspective. Teams with more control over the solution process and solution implementation (*Team Autonomy*) also reported greater gains in *Task KSA*. Furthermore, this finding is consistent with previous findings from the team literature, which indicate that increased team autonomy is associated with increased team learning (Cohen & Ledford, 1994; Kirkman & Shapiro, 1997; Hyatt & Ruddy, 1997; Kirkman & Rosen, 1999; Edmondson, 2002). Thus, if Kaizen events are intended to be used as a continuous improvement training tool, this finding would suggest that *Team Autonomy* should be carefully preserved.

Internal Processes had a significant direct relationship to *Task KSA*. This was the only direct relationship shared by the two social system outcome variables (*Attitude* and *Task KSA*); however, both social system variables had a common indirect relationship (i.e., *Goal Clarity*). While not conceptually a necessary condition, harmonious internal workings would likely promote gains in *Task KSA*. In addition, *Internal Processes* also measures the extent to which there was discussion of ideas within the team. Discussing ideas and developed shared meanings is a fundamental process of group learning (Crossan et al., 1999; Edmondson, 2002). One additional variable – *Goal Clarity* – acted indirectly on *Task KSA* through *Internal Processes*. The relationship of *Internal Processes* to *Goal Clarity* has already been discussed in the previous section.

A second event process variable, *Affective Commitment to Change*, also had a positive direct relationship to *Task KSA*. This finding indicates that events with higher initial commitment ("buy-in") to the event objectives (measured at the Kickoff meeting) also had higher gains in *Task KSA*. These findings are particularly interesting since *Affective Commitment to Change* showed no relationship to *Impact on Area* or either of the other two technical system measures. Although not differentiated in the research model (Figure 1), it would seem particularly likely that this variable would be related to implementation (i.e., *Impact on Area* or *% of Goals Met*), effects that were not observed in this research. The exact nature of the relationship between *Affective Commitment to Change* and *Task KSA* is not known. Also, as mentioned previously, in any observational study, correlation cannot be assumed to indicate causation, although the fact that *Affective Commitment to Change* was measured before *Task KSA* strengthens the conceptual argument for causation. It may be that teams with a higher initial buy-in work harder – i.e., are more dedicated and diligent – at applying the tools (e.g., Keating et al., 1999), thus resulting in greater gains in *Task KSA*. Although team effort was not measured in the current research, future research could include measures of perceived team effort to determine whether team effort appears to mediate the relationship between *Affective Commitment to Change* and *Task KSA*, as proposed above. Another potential explanation for the relationship between *Affective Commitment to Change* and *Task KSA* is that when team members believe the event itself is worthwhile (i.e., high *Affective Commitment to Change*), they may be more likely to feel they have gained worthwhile KSAs from the event.

5.4 Significant Predictors of Impact on Area

Figure 4 depicts the overall model of significant relationships between *Impact on Area* and the predictor variables which were identified in this research. Table 56 contains the relative effect sizes – i.e., GEE regression coefficients – for the variables in the final model.

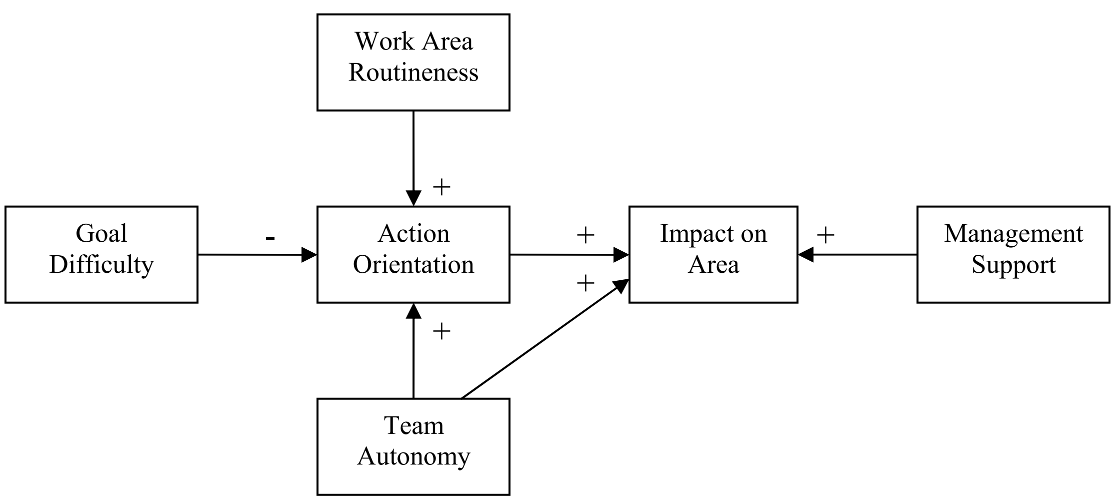

Figure 4. Overall Model for Significant Predictors of Impact on Area

Table 56. Effect Size Table for Impact on Area

Predictor	Direct Effect	Indirect Effect
Action Orientation	.243	
Team Autonomy	.342	.182
Management Support	.262	
Goal Difficulty		-.170
Work Area Routineness		.106

This research found that the most significant predictors of team member perceptions of the impact of the given event on the target system (*Impact of Area*) were *Management Support* (direct positive), *Action Orientation* (direct positive), *Team Autonomy* (direct and indirect positive), *Goal Difficulty* (indirect negative) and *Work Area Routineness* (indirect positive). *Management Support*, *Action Orientation* and *Team Autonomy* demonstrated a positive relationship with *Impact on Area*, while *Goal Difficulty* and *Work Area Routineness* demonstrated a negative relationship with *Impact on Area*.

Action Orientation was the only event process variable with a significant relationship to *Impact on Area*. *Action Orientation* describes team member perceptions of the relative amount of time their team spent in the target work area versus "offline" in meeting rooms. For the events studied in this research, increasing levels of *Action Orientation* were associated with increased perceptions of *Impact on Area*. This lends support to the hypothesis that increased levels of *Action Orientation* denote increased focus on implementation – i.e., making changes to the target work area. This is also supported by analysis of the contextual data provided in the team activities log. Although none of the reported levels of *Impact on Area* in this research were low per se in terms of the survey response scale – it would thus appear that, in non-implementation events, team members likely based their perceptions on the projected future impact of the event on the target system if their solution were fully implemented – increased *Action Orientation* was associated with increased perceptions of *Impact on Area*. Increased focus on "hands on" activities (e.g., implementation) has been suggested as one of the main features that differentiate Kaizen events from more traditional continuous process improvement activities – e.g., quality circles and the continuous process improvement teams used in Total Quality Management – where the end result of team activities is often an action plan for change which is then presented to management for approval (Mohr & Mohr, 1983; Cohen & Bailey, 1997; Laraia et al., 1999). In these traditional activities, there is generally more focus on analysis and less focus on experimentation – i.e., testing out solutions right away. In addition, in traditional continuous process improvement activities, in many cases, team solutions may not be implemented until weeks or months after the presentation to management, if they are implemented at all. The lack of immediate, apparent impact of their efforts, as well as the fact that management could ultimately overrule and fail to implement team solutions, have been suggested as reasons why many continuous process improvement programs have failed to be sustained. The lack of immediate short-term returns on investment is hypothesized to have resulted in lower employee buy-in to the programs and ultimate failure to sustain the programs (e.g., Lawler and Morhman, 1985, 1987; Keating et al., 1999).

Again, the Kaizen event facilitator would seem to play a key role in enabling *Action Orientation*, both through managing key input variables – to be discussed more presently – and through his or her role in directly assisting in the coordination of team activities. By establishing ground rules – e.g., the importance of "hands-on" experimentation or "trystorming;" "better" vs. "perfect," etc. (Mika, 2002; Farris & Van Aken, 2005) – and by encouraging the team to spend its time in the target work area versus the meeting room, the facilitator could directly influence the extent of team *Action Orientation* and thereby potentially ultimately increase *Impact on Area*. This research suggests that key ways that the facilitator or others in the organization could indirectly influence *Impact on Area* through *Action Orientation* appear to include event planning activities related to event scoping and boundary control. Developing team objectives of appropriate difficulty (*Goal Difficulty*), selecting target system of appropriate complexity (*Work Area Routineness*), and establishing *Team Autonomy*.

Goal Difficulty was negatively associated with *Action Orientation*. Teams with higher levels of *Goal Difficulty* believed that they spent relatively more time in their meeting room versus in the target work area (i.e. lower *Action Orientation*), compared with teams reporting lower levels of *Goal Difficulty*. As described above, it is likely, although as yet untested, that lower levels of *Action Orientation* are associated with more time spent in analysis activities – i.e., understanding the problem and the target system and designing a solution – versus implementation activities – i.e., implementing and testing solutions. The finding that *Goal Difficulty* appears to have a negative relationship to *Impact on Area* through *Action Orientation* appears to contradict the common recommendation in the Kaizen event practitioner literature that team goals should be challenging "stretch" goals (e.g.. LeBlanc, 1999; Minton, 1998; Rusiniak, 1996; Cuscela, 1998; Bradley & Willett, 2004; Bicheno, 2001; Tanner & Roncarti, 1994; Treece, 1993; Kumar & Harms, 2004; Gregory, 2003). However, as has already been mentioned, although *Goal Difficulty* has a negative relationship to *Impact on Area* (and *% of Goals Met*), it has a positive relationship to *Task KSA*, suggesting a trade-off between employee learning and technical system impact. This finding is also somewhat surprising given that research on goal setting has fairly consistently reported a positive relationship between *Goal*

Difficulty and task performance (e.g., Locke & Latham, 1990). However, these goal setting studies have generally centered on the relationship between "objective" (i.e., researcher-assessed) levels of *Goal Difficulty* and task performance in controlled experiments where extensive training on the task is provided to subjects prior to experimental manipulation. On the other hand, Martin, Snell and Callahan (1999) found a negative relationship between subjective (i.e., respondent-perceived) *Goal Difficulty* and task performance for goals that were also "objectively" difficult – i.e., as assessed by the researcher. Similarly, Earley, Connolly & Ekegren (1989) found that, when training on the specific tactics and strategies needed to achieve the goals was omitted, subjects with hard externally assigned goals had lower task performance than subjects that were merely told to "do their best."

In general, the findings of this research suggest that, in the context of Kaizen events, overly difficult goals – i.e., goals that are perceived by team members as not achievable (Schneiderman, 1988) – may be associated with lower outcomes. This agrees with suggestions – albeit the minority voice – in the Kaizen event practitioner literature that Kaizen events should avoid tackling problems that are "too big" in scope (i.e., difficult) (Rusiniak, 1996; Sheridan, 1997b; "Get Smart, Get Lean," 2003; Gregory, 2003), as well as problems that require advanced statistical analysis or other complex problem-solving tools (Bradley & Willett, 2004; Harvey, 2004), which would also likely be reflected in increased perceptions of *Goal Difficulty*. These variables – i.e., event scope and the complexity of the needed analysis methods – could be studied in future research. This finding is not, however, interpreted to imply that the absolute magnitude of improvement targeted by an event should not be large, but rather that the event goals should be scoped such that team members perceive them as attainable. For instance, if there is a lot of "low hanging fruit" – i.e., obvious inefficiency -- in a given work area, goals targeting a relatively large magnitude of improvement could be perceived as relatively less difficult compared to the same goals set for a more mature work area. In addition, Martin, Snell and Callahan (1999) suggests that perceived *Goal Difficulty* can be directly impacted by the mechanisms used to communicate the goals. Through training and team discussions in the Kickoff meeting, facilitators may be able to help team members perceive objectively difficult goals as more achievable – i.e., reducing perceived *Goal Difficulty* – thus potentially ultimately increasing *Impact on Area*. Furthermore, additional countermeasures appear to be available by manipulating other variables related to *Impact on Area*. For instance, increasing *Team Autonomy* could act as a countermeasure for *Goal Difficulty*, both by increasing *Action Orientation* to compensate for the negative impact of *Goal Difficulty* – note, the measurement scales and indirect effect sizes are similar for *Goal Difficulty* and *Team Autonomy* – and by directly influencing *Impact on Area*. Or, *Management*

Support could be increased to directly compensate for the effect of *Goal Difficulty*. Note, the effect size of *Management Support* is larger than the effect size of *Goal Difficulty*, and the measurement scales are the same.

The finding that *Goal Difficulty* is negatively related to *Impact on Area* is therefore not interpreted to suggest that Kaizen event facilitators or other organizational personnel should never use a Kaizen event to attack a clearly difficult problem, but rather, this finding suggests that they should recognize the apparent trade-off between *Goal Difficulty* and perceived *Impact on Area* and avoid setting difficult goals simply for the sake of "challenging" the team, as recommended in the Kaizen event practitioner literature. In addition, this finding also suggests that organizations should avoid the "hammer-nail" syndrome – i.e., the temptation to use Kaizen events to address every type of organizational problem. For problems that require more complex analysis methods and/or occur in larger scope processes – i.e., cutting across several functional boundaries – other problem-solving mechanisms, such as Six Sigma teams, continuous process improvement teams or hybrid Kaizen event-continuous process improvement teams (Farris & Van Aken, 2005), may be more appropriate.

Similarly to *Goal Difficulty*, it appears that *Work Area Routineness* may have a positive relationship to *Impact on Area* through *Action Orientation* – i.e., the study findings are consistent with the mediation hypothesis. Events in more complex (less routine) work areas also demonstrated lower *Action Orientation* and, in turn, lower perceptions of *Impact on Area*, presuming the hypothesized direction of causality in the mediation model is correct. This finding has strong face validity, because it seems likely that the team may need to spend more event time "offline" during events that occur in more complex target work areas, both since designing the solution make take more time due to greater process complexity and since the process may not be directly observable – e.g., a knowledge process and/or a process that cuts across cross-functional boundaries. This also agrees with propositions in the Kaizen event practitioner literature that the work areas to be targeted by Kaizen events should be "reliable," with relatively predictable flow patterns and outputs – i.e., some baseline levels of standardization (LeBlanc, 1999; Bradley & Willett, 2004). This also agrees with the general proposition in the lean manufacturing literature that "traditional" lean practices – often, the same as those associated with Kaizen events – are most effective after some baseline standard of process reliability has been established – i.e., in more routine processes (e.g., James-Moore & Gibbons, 1997; Cooney, 2002: Papadopoulou & Ozbayrak, 2005), as well as the focus in the Kaizen and Total Quality Management literature on establishing process reliability as a necessary baseline to additional quality improvements (e.g., Imai, 1986). However, the lean manufacturing literature contains many recommendations of how lean principles can be adapted for more variable work environments, although such adaptation remains challenging (for instance see Roby, 1995; Jina et al., 1997; Bowen & Youndahl, 1998; Bozzone, 2002 and Murman et al., 2002). This finding may therefore suggest that, if the target process lacks baseline reliability (capability), some other improvement activities be targeted at improving the standardization of the target work areas prior to conducting Kaizen events. However, this finding is not interpreted to suggest that, in general, companies should never hold Kaizen events in complex work areas, since, it is possible, even likely, that more complex work areas – particularly those in knowledge processes, such as engineering – may be the overall bottlenecks in organizational improvement (Murman et al., 2002). In addition, countermeasures may likely be employed to offset the loss in *Action Orientation*, which is the variable which appears to be directly affected by *Work Area Routineness*. For instance, the facilitator or team leader could structure the event such that more time is spent observing the work area or in "hands-on" experimentation (simulation, "trystorming," etc.). Or, the facilitator could attempt to compensate through one of the other variables believed to affect *Impact on Area* through *Action Orientation* – *Goal Difficulty* or *Team Autonomy*. In addition, facilitators could increase *Team Autonomy* or *Management Support* to directly compensate for the reduction in *Impact on Area*.

Team Autonomy had both an apparent indirect relationship to *Impact on Area* (through *Action Orientation*) and a direct relationship. Provided the hypothesized direction of causality in the mediation model is correct, it would appear that increasing *Team Autonomy* regarding what solution is developed, the process used to develop the solution, and ability to implement the solution contributes to higher levels of *Action Orientation*, which, in turn contribute to increased perceptions of *Impact on Area* (see discussion above of the ways *Action Orientation* is hypothesized to effect *Impact on Area*). The finding that *Team Autonomy* also appears to have a direct relationship to *Impact on Area* could suggest that teams that are given a greater amount of freedom in developing solutions believe that they developed better solutions – i.e., solutions with greater *Impact on Area* – versus teams given less freedom in developing solutions. This finding could also suggest that teams with greater *Team Autonomy* are more likely to be able to actually implement their solutions during, leading to increased levels of *Impact on Area* over teams that were not given as much freedom to implement.

The final variable which appears to have a significant direct relationship to *Impact on Area* is *Management Support*. Although this variable was not significant at the 0.10/p level in the final regression model, there appears to be sufficient evidence to suggest a likely relationship (i.e., $p < 0.05$). It is perhaps surprising that *Management*

Support does not appear to act indirectly through *Action Orientation*. Given the fact that *Management Support* appears to be directly versus indirectly related to *Impact on Area*, one potential hypothesis for the nature of this relationship is that, by providing the necessary tools and access to others in the organization, *Management Support* can assist teams with high *Action Orientation* and *Team Autonomy* to actually implement their solutions, in turn increasing *Impact on Area* versus teams with lower *Management Support* and higher levels of the other two variables. The findings of this research suggest that *Management Support* is a high leverage variable for organizations since it has a positive relationship to three out of five outcomes (the only outcome variables without a direct or indirect relationship to *Management Support* are *Task KSA* and *% of Goals Met*). This agrees with previous team research which has found *Management Support* to be a significant predictor of technical system and social system team outcomes (e.g. Campion et al. 1993; Hyatt & Ruddy, 1997; Doolen et al., 2003a). However, despite the rarely questioned theoretical importance of *Management Support*, Campion, Medker & Higgs (1993) noted that *Management Support* traditionally received less empirical attention in team research than other variables believed to impact team effectiveness. The results of this research reinforce the need for researchers to explicitly measure *Management Support* in studies of team effectiveness and for organizations to carefully plan and consistently deliver high levels of *Management Support*.

The *Management Support* attributes specifically included in the final version of the *Management Support* variable used in the analysis include sufficiency of resource support – i.e., materials and supplies, and equipment -- and sufficiency of access – i.e., help from others in the organizations. Fortunately for organizations, these aspects of *Management Support* appear to be fairly easy for managers to control. For instance, organizational policy can be adapted to help ensure that teams will have access to needed subject matter experts and support resources (e.g., maintenance) during the event, either as members of the team or "offline" support personnel. The strategy of making sure that outside subject matter experts are available during the event to help the team has been recommended in the Kaizen event practitioner literature (McNichols et al., 1999; Martin, 2004; Sheridan, 1997b; Bradley & Willett, 2004; Bicheno, 2001; Adams et al., 1997; Wittenberg, 1994; Tanner & Roncarti, 1994; Gregory, 2003; Taylor & Ramsey, 1993). In terms of resource support, in addition to standard resource checklists for event planning (e.g., Mika, 2002; Vitalo et al., 2003), one innovative suggestion is the use of a standard Kaizen event equipment cart containing commonly needed tools and supplies, which serves as a mobile office for the Kaizen event team during the event (Taylor & Ramsey, 1993). Since *Management Support* was designed to represent a global variable – and not just the three specific aspects mentioned above – it is also hypothesized that increasing other aspects of *Management Support* – i.e., encouragement from top management, etc. – could also facilitate increased *Impact on Area*, provided the direction of causality in the research model is correct.

5.5 Significant Predictors of Overall Perceived Success

Figure 5 depicts the overall model of significant relationships between *Overall Perceived Success* and the predictor variables which were identified in this research. Table 57 contains the relative effect sizes – i.e., GEE regression coefficients – for the variables in the final model.

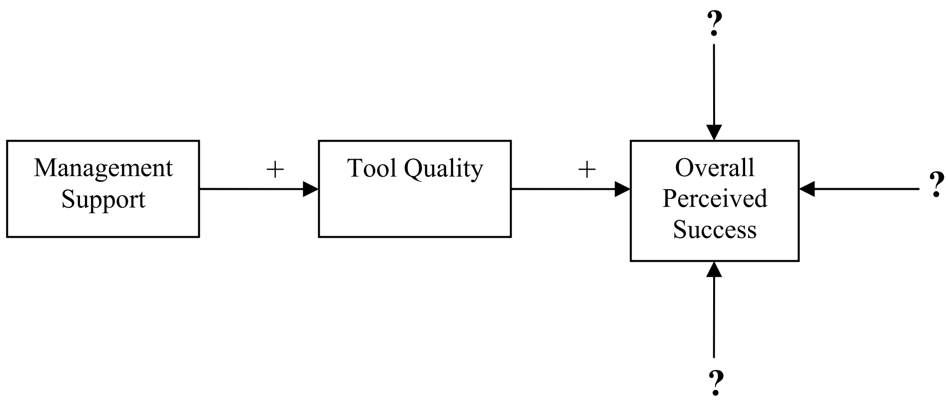

Figure 5. Overall Model for Significant Predictors of Overall Perceived Success

Table 57. Effect Size Table for Overall Perceived Success

Predictor	Direct Effect	Indirect Effect
Tool Quality	.621	
Management Support		.316

On the whole, this research revealed the need for additional, future research on the predictors of facilitator perceptions of overall event success. Only one factor studied in this research was a significant direct predictor of *Overall Perceived Success*. One additional variable (*Management Support*) was a nearly significant indirect predictor. In addition, the final regression model of *Overall Perceived Success* only accounted for about 12% of the total variation in *Overall Perceived Success*, indicating that the vast majority of the variation in this measure cannot be explained by the variables studied in this research. Even the maximum adjusted R^2, which was for a regression model containing some non-significant predictors, was around 14% (using OLS estimates) and the maximum R^2 – i.e., for the regression model containing all 14 candidate predictors – was only around 27% (again, using OLS estimates).

These results are somewhat surprising, given the fact that the event input and event process factors studied in this research were chosen because they were hypothesized to be the factors most closely related to overall Kaizen event success, as suggested in the Kaizen event practitioner literature and related organizational literature (see Chapter 2). However, it should be noted that several aspects of the outcome measure itself could have limited the ability of the research to find results.

First, the variation in the *Overall Perceived Success* measure was relatively low for the events studied in this research. This is likely also related to the fact that *Overall Perceived Success* is a single item measure, which will be discussed more presently. All but three of the events studied received a score of 4.0 or above on the six-point rating scale. Thus, although the distribution of *Overall Perceived Success* was relatively symmetric and *Overall Perceived Success* was measured on an interval scale, it suffered from range restriction – i.e., the vast majority of events received a 4.0, 5.0 or 6.0 rating. This range restriction makes *Overall Perceived Success* less suitable for regression, since it is less likely that a true, linear relationship will exist between the predictors and the outcome variable (Field, 2005). Although this research treated *Overall Perceived Success* as a continuous variable, because it is generally desirable to avoid transforming a continuous variable into an ordinal or categorical variable, in future, follow-up research using this data set, it is possible that multinomial regression could be applied – i.e., treating each of the response intervals as categories, with the lowest three responses likely excluded or grouped in one category. However, SAS PROC GENMOD does not currently have the capability to perform multinomial regression with GEE unless the independence correlation structure is assumed. Although assuming an independence correlation structure might provide an adequate fit given the characteristics of the data –i.e., since it appears that there is relatively more variation within versus across groups – it would not be consistent with the assumption throughout the rest of this research of correlation between teams within organizations. Thus, the suggestion is that this regression might be performed as a post-hoc analysis but would not be considered as part of the primary analysis. Also, as discussed below, expanding *Overall Perceived Success* into a multi-item measure could help increase the variance of this variable, which might make it a better fit for Gaussian regression modeling in future research.

The *Overall Perceived Success* variable was a single item measure, rather than a multi-item survey scale. However, the current research suggests that *Overall Perceived Success* may actually be a multidimensional construct. Factor analysis of the outcome variables in the Report Out Survey (see Chapter 3), which contained an identical *Overall Perceived Success* measure, indicated that this item loaded onto both a technical scale (*Impact on*

Area) and a social system scale (*Attitude*). Thus, although it is not certain that this relationship will hold for facilitator perceptions as well as team perceptions, perhaps the *Overall Perceived Success* variable could be expanded into a multi-item measure containing both technical and social system elements – i.e., perceptions of impact on area, perceptions of top management satisfaction, etc. Or, separate scales could be used to measure perceptions of technical impact – i.e., the *Impact on Area* scale could be used to gather data from the facilitator, as well as team members – and social system outcomes – i.e., a new scale could be developed to measure facilitator perception of top management satisfaction, as well as the satisfaction of other key stakeholders.

Research on the predictors of *Overall Perceived Success* appears to be particularly important since facilitator comments in the Event Information Sheets indicate that facilitator ratings of *Overall Perceived Success* may be related to top management satisfaction with event results (for instance, see Farris et al., 2006) and the fact that *Overall Perceived Success* was not significantly correlated with goal achievement (*% of Goals Met*). In particular, there does not appear to be a strong relationship between *% of Goals Met* and *Overall Perceived Success* for the small number of events that were considered unsuccessful by their facilitators. The two events that received a rating of 1.0 ("strongly disagree") in the current research achieved 91% and 100% of the goals, respectively. The other event receiving an *Overall Perceived Success* rating less than 4.0 (a 2.0) achieved 83% of the goals, which is higher than many events receiving a 4.0, 5.0 or 6.0 rating. It is not immediately apparent which additional event input and event process variables should be studied in future research. One way to shed light on this may be to conduct interviews with event facilitators to identify what variables they believe are mostly strongly related to overall event success. Then, these additional variables and the predictors found to be significant in this research – i.e., *Tool Quality* and *Management Support* – could be studied in future research involving a second set of facilitators to avoid same source bias.

Tool Quality was the only variable that appears to have a significant direct relationship to *Overall Perceived Success*. It may be that teams that more effectively use their problem-solving tools (i.e., have higher *Tool Quality*) are ultimately more successful overall, as measured by higher levels of *Overall Perceived Success*. This is the relationship that has been hypothesized in this research, where *Tool Quality* has been treated as a process variable. Conversely, it could be that facilitators base their rating of *Overall Perceived Success* at least in part on the quality of the team's tool use – i.e., that *Tool Quality* can be considered a component of *Overall Perceived Success*, and could therefore be treated as an item in a revised *Overall Perceived Success* scale. While this second alternative seems less likely, this proposition could be further evaluated in future research involving a multi-item measure of *Overall Perceived Success* and perhaps a *Tool Quality* rating from team members, as well as the facilitator.

As mentioned, there is marginal support that *Management Support* may have a significant indirect effect on *Overall Perceived Success* through *Tool Quality*, again, presuming the hypothesized direction of causality in the mediation is correct. This result appears to reinforce the key role of *Management Support*, since it is a direct or indirect predictor of three out of five outcomes. In addition, these results have strong face validity, since *Management Support* – i.e., having needed materials, equipment/supplies, and assistance from other organizational personnel – would seem necessary for enabling effective tool use (i.e., higher *Tool Quality*).

Additional post-hoc analyses were performed to determine whether *Overall Perceived Success* was a derived outcome measure – i.e., based on the other four outcome measures – *Attitude*, *Task KSA*, *Impact on Area* and *% of Goals Met* – instead of a direct outcome measure. However, no significant predictors were found in the regression of *Overall Perceived Success* on these other four outcome measures. In addition, a backward selection analysis was used to determine whether, if *Tool Quality* were withheld from the analysis, additional predictors of *Overall Perceived Success* would emerge. However, no other significant predictors were found in this analysis. All these results support the need for additional research to understand the predictors of *Overall Perceived Success*.

5.6 Significant Predictors of % of Goals Met

Figures 6 and 7 depict the overall models of significant relationships between *% of Goals Met* and the predictor variables, which were identified in this research for the continuous representation and dichotomous representation, respectively. Tables 58 and 59 contains the relative effect sizes – i.e., GEE regression coefficients – for the variables in the final model for the continuous representation and dichotomous representation, respectively.

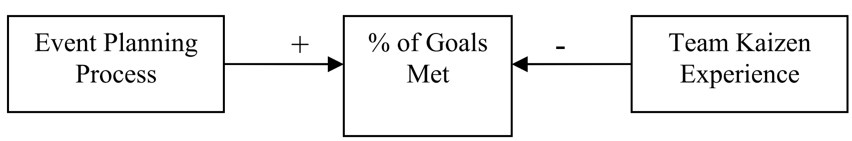

Figure 6. Overall Model for Significant Predictors of % of Goals Met (Continuous Variable)

Table 58. Effect Size Table for % of Goals Met (Continuous Variable)

Predictor	Direct Effect
Team Kaizen Experience	-.486
Event Planning Process	.149

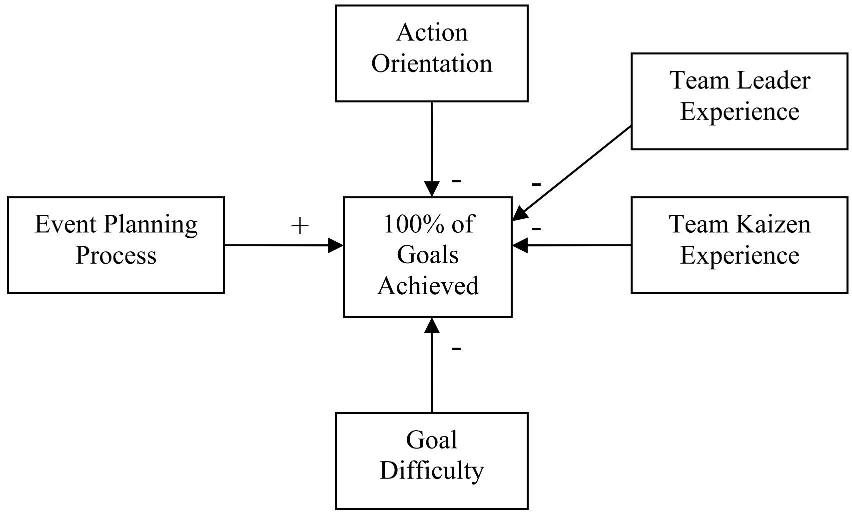

Figure 7. Overall Model for Significant Predictors of Goal Achievement (Dichotomous Variable)

Table 59. Effect Size Table for % of Goals Met (Dichotomous Variable)

Predictor	Direct Effect
Goal Difficulty	-1.979
Team Kaizen Experience	-4.254
Team Leader Experience	-4.262
Event Planning Process	4.554
Action Orientation	-1.901

For the continuous regression, there were two significant predictors of *% of Goals Met*: *Team Kaizen Experience* and *Event Planning Process* (hours planning). *Event Planning Process* – i.e., the number of hours spent planning the event – had a positive relationship to *% of Goals Met*. This suggests that the amount of time the facilitator or others in the organization are allowed to spend planning the event does have a substantial impact on event performance. Organizations appear to face a likely trade-off in event performance when they hold "last minute" events (see for example Letens et al., 2006) or attempt to squeeze too many events in over too short a timeframe, thereby reducing the amount of time allowed to plan each event. In the current research, the median amount of time spent planning the event was 6.0 hours (i.e., three quarters of an eight-hour work day). However, the amount of time planning ranged from 0.5 hours to 120 hours (i.e., three 40-hour work weeks). In fact, due to this large range of variation and outlying observations on the high end, *Event Planning Process* was log transformed to reduce the variation prior to the regression analysis.

Based on the events studied, this research would seem to suggest two things related to the *Event Planning Process*. First, in order to allow a substantial amount of time to be devoted to event planning, it would appear to be beneficial to adopt the practice of having one or more full-time event facilitators or coordinators –i.e., one or more individuals who spend the majority of their time planning and conducting events (Heard, 1997; "Keys to Success," 1997; Bicheno, 2001; Foreman & Vargas, 1999). However, these results also suggest that organizations must resist the urge of overburdening these fulltime personnel with a larger number of events than they can effectively plan, conduct and follow-up. The Kaizen event practitioner literature is full of examples of organizations that have become so convinced of the effectiveness of the Kaizen event process that they run one or more events per week (e.g., Vasilash, 1997; Sheridan, 1997b; LeBlanc, 1999). In fact, in the organizations studied in this research, most conducted three or more events per facilitator per month. In these cases, based on the high volume of events conducted, with little or no downtime in between, one wonders how the organizations are able to effectively plan events and to devote resources to follow-up – and, as this research indicates, perhaps they do not do this as effectively as they might, had they allowed more time for planning activities. Based on the results of this research, setting an organizational objective of each facilitator running an event each week – or every three out of four weeks – could likely be too high a workload to allow adequate planning for the events, not to mention adequate devotion of resources to follow-up. This research suggests that organizations could be well served – both in terms of initial results and results sustainability – by spacing out events to one or two per facilitator per month, to allow adequate event planning and the ability to devote resources to follow-up.

Team Kaizen Experience had a direct negative relationship to *% of Goals Met*. This effect is surprising and warrants further investigation in future research. It could be that the actual magnitude of *Team Kaizen Experience* contributes directly to lower success. For instance, it is possible that more experienced teams enter the event with more preconceived notions and thereby develop less creative solutions, with the ultimate result of lower goal achievement (i.e., *% of Goals Met*). However, it could also be that *Team Kaizen Experience* is a correlate for some other, unmeasured predictor variable. One likely correlate would be event complexity or overall event scope. *Goal Difficulty* was measured in this research. However, since *Goal Difficulty* was measured through a perceptual survey measure, it is necessarily at least somewhat relative to the experience (i.e., perceptions) of team members. Therefore, more experienced team members might perceive the same level of *Goal Difficulty* for an event that is more objectively challenging than less experienced team members might perceive for an event that is less objectively challenging. In support of this proposition, a post-hoc analysis of the correlation between *Team Kaizen Event Experience* and *Goal Difficulty* found no significant correlation between the two measures at the team level. In addition, another post-hoc analysis reported in Chapter 4 found no significant relationship between two measures of complexity (*Number of Main Goals* and total *Number of Goals*) and *% of Goals Met*. However, to further investigate the proposition that *Team Kaizen Experience* may related to overall event scope, in future research, event scope could be more formally measured, either through perceptual measures or through designed – rather than post hoc – objective measures. In addition, perhaps in future research the facilitator, as well as the team, could be asked to rate the difficulty of the team's goals, thus providing a consistent baseline by an internal expert. Just as team members are asked to rate *Goal Difficulty* at the Kickoff Meeting, before the measure can be influenced by event problem-solving activities, the facilitator could also be asked to rate the goals prior to the event, to avoid any contamination by the problem-solving activities of the event.

In the dichotomous variable analysis – i.e., the analysis to discover what characteristics can be used to predict events that were 100% successful in achieving their goals versus events that were less than 100% successful – there were three additional predictors of Goal Achievement: *Goal Difficulty*, *Team Leader Experience* and *Action Orientation*.

As hypothesized, *Goal Difficulty* was negatively related to Goal Achievement – i.e., the probability that events were 100% successful in achieving those goals. Events that were rated by team members as having more difficult goals did have lower levels of Goal Achievement. This may suggest that team members are fairly good judges of the difficulty of the goals relative to their abilities. It may also suggest that, in events where *Goal Difficulty* was rated more highly, team members were not given adequate training on the strategies needed to achieve their goals.

Contrary to hypothesis, *Team Leader Experience* was also negatively related to Goal Achievement. All the recommendations in the team literature, project management literature and the Kaizen event practitioner literature suggest that having a more experienced team leader should increase the likelihood of event success (see Chapter 2). However, as described earlier in this chapter, it is possible that increased *Team Leader Experience* is associated with greater direction of the team activities by the team leader – i.e., a stronger, more controlling or directive leader – and lower contribution of additional members to the overall strategy of the event solution process. This effect would be consistent with team research which has found that, at least in certain task settings, more controlling (i.e., directive) leaders are associated with lower team performance in terms of goal achievement (e.g., Durham et al., 1997). Conversely, this finding could indicate an unmeasured variable that is correlated with both *Team Leader Experience* and *% of Goals Met*. One possible correlate is event complexity (see the discussion about *Team Kaizen*

Experience). Similarly to *Team Kaizen Experience*, future research is warranted to discover more about the relationship between *Team Leader Experience* and *% of Goals Met*.

Another interesting finding was that *Action Orientation* was also negatively related to Goal Achievement. While *Action Orientation* has been cited in the practitioner literature as a defining characteristic of Kaizen events – and a mechanism that promotes implementation and experimentation -- it is not clear, even from the practitioner resources to what extent *Action Orientation* is suggested to be directly related to increased goal achievement, particularly given other event characteristics (e.g., *Goal Difficulty*, etc.). While many practitioner resources would suggest that teams spend as much of their time in the target work area as possible (e.g., LeBlanc, 1999; Redding, 1996; Smith, 2003; Martin, 2004; Sheridan, 1997b; Patton, 1997; Vasilash, 1993; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998; Sabatini, 2000; Tanner & Roncarti, 1994; Larson, 1998a; Treece, 1993; Taylor & Ramsey, 1993; Foreman & Vargas, 1999), these results suggest that, at least in some cases, teams might be better served by spending a substantial portion of time conducting training, establishing ground rules and analyzing the problem – activities that are generally conducted "offline" in meeting rooms versus on the shop floor – i.e., in the target work area. There is some evidence from this research that laying more groundwork and allowing more time to developing, versus implementing, solutions may ultimately lead to increased goal achievement, particularly given some event designs. As demonstrated by the mediation relationship between *Action Orientation* and *Goal Difficulty* for *Impact on Area*, *Action Orientation* and *Goal Difficulty* are significantly, negatively correlated. The negative relationship between *Goal Difficulty* and *% of Goals Met* has already been discussed. However, the fact that *Action Orientation* had an additional, significant negative effect may suggest that, when events with more difficult goals adopted a relatively high degree of *Action Orientation*, a further negative effect resulted. Or, similarly to *Team Kaizen Experience* and *Team Leader Experience*, it is possible that *Action Orientation* is correlated with an unmeasured predictor variable such as event scope. Thus, future research that includes direct measures of event scope is also of interest for *Action Orientation*.

5.7 Limitations of the Present Research

Section 1.8 originally presented some of the limitations of the current research – specifically the limited number of variables and organizations studied. While these limitations are not discussed in detail here (due to the detailed treatment given in Chapter 1), it is worth noting here that additional event input, event process and event outcome variables could have been studied. Thus, the findings of this research only hold for the variables studied in this research and do not rule out significant effects from unstudied variables. Related to this observation, in all the regression models, a substantial amount of variation (i.e., at least 40%) remained unexplained, thus indicating the likely presence of other explanatory variables that were not measured. However, the current set of variables were chosen for theoretical reasons – i.e., the fact that the Kaizen event practitioner literature and literature on related organizational mechanisms (i.e., projects and teams) suggested that they were the most theoretically likely predictors. Chapter 6 further describes the opportunity of examining additional variables in future research.

Similarly, the boundary criteria applied to select organizations – i.e., manufacturing organizations that are relatively experienced in using Kaizen events and hold Kaizen events frequently – as well as the non-random nature of sampling at the organizational level, mean that the results of this research may not hold for organizations of markedly different characteristics. Conversely, however, there was quite a bit of variety across organizations in terms of the types of problems and work areas targeted, including events in non-manufacturing processes. In addition, the boundary selection criterion of relatively "high" experience in using Kaizen events means that these organizations are more likely to be "best practice" organizations. They have continued to use Kaizen events, because presumably, they have found them to be effective in achieving organizational objectives – although all Kaizen event coordinators in the participating organizations indicated that there was some variation in outcomes across events, which is also indicating by the negative intraclass correlation values in the GEE regression models. Finally, as has been mentioned, the sample size at the organizational level was relatively small (i.e., six organizations), and, as will be discussed in Chapter 6, future research involving more organizations is desirable to test the robustness of these results and to allow the investigation of organization-level variables.

Another limitation is that this research only studied the initial outcomes of Kaizen events. No indication of the level of sustainability of these outcomes – or the mechanisms related to sustainability – is therefore provided. As Chapter 6 indicates, this is an area where future research is needed and is also a focus area within the larger VT-OSU study of Kaizen events.

One additional limitation of this research was discussed at the beginning of this chapter. As an observational field study, this research lacks experimental control and the findings are based on methodological and theoretical arguments for causality – i.e., timing of measurements, nature of measurements, related organizational theory, etc. – rather than empirical proof of causality through experimental control. This is an issue which has been long recognized as a limitation of non-experimental field research; however, such theoretical observational research still remains valuable for building and testing theory, particularly in cases of complex real world phenomena and particularly for exploratory work, such as this research. Controlled laboratory experiments involving complex business processes often lack generalizability due to the non-random nature of the subject pool – i.e., often college students – and the artificial and generally overly simplified environment induced in the experimental setting. Quasi-experiments in the field often have increased validity in these areas but lack precise control. There are always contextual factors – measured or unmeasured – that can impact results, which is a similar problem to that faced in observational studies; however, there is a stronger argument for the direction of causality in quasi-experiments versus observational studies. Using a quasi-experimental design prior to this study did not appear feasible. First, since before this study, none of the design suggestions in the Kaizen event practitioner literature had been empirically investigated, there were a large number of potential effects to be tested, which did not prove to be problematic in the current study design but would have been more problematic in an experimental design due to the increased resource investment for organizations. This research provides some groundwork for future quasi-experimentation by narrowing down the hypothesized sets of relations for each outcome variable to parsimonious sets for further testing. Second, the measures used in this study had not yet been validated – i.e., through factor analysis, etc. – thus representing an increased risk for organizations in terms of time and resources invested if the measures proved unreliable. Third, related to the first point, since the effects hypothesized in Kaizen event literature had not been tested, the risk for participating organizations in adopting these suggestions could have been large – i.e., in terms of decreased event outcomes. Since these design suggestions have now been analyzed to indicate which variable demonstrate positive relationships with outcomes, perhaps a quasi-experimental design could be adopted in future research where a baseline of organizational performance is taken on several events before and after adopting the design suggestions associated with positive outcomes in this research.

Another limitation, which is related to the small sample size at the organizational level and has also been discussed in Chapter 4, is that the small sample size at the organizational level precluded the ability to test for differences in regression slopes across organizations. Aggregate differences – i.e., differences in regression intercepts – were accounted for through the use of GEE; however, as discussed in Chapter 4, GEE does not allow the separate modeling of differences in regression slopes across organizations. As discussed in Chapter 6, a larger sample size at the organizational level would allow the investigation of differences in slopes across organizations as well as organization-level variables through HLM.

The final limitation noted in this section is the exploratory nature of this research. Due to the lack of prior research to establish the order or importance of the study variables, the regression models in this research were built through an exploratory variable selection process. Although several different approaches were used to attempt to establish convergence, there is always a chance that the "best" model – i.e., the one that most appropriately represents the relationship between the outcome variables and the input variables – was not the final model selection. This risk is always at least somewhat present in regression analysis, but particularly in exploratory work.

CHAPTER 6: CONCLUSIONS

This chapter summarizes the overall findings of this research and describes areas of future research that were identified as a result of this study. There are four primary areas of future research identified: 1) additional testing of model robustness with an increased sample size at the organizational level; 2) testing of additional model parameters; and 3) additional research on the sustainability of event outcomes.

6.1 Summary of Research Findings

Table 60 provides a summary of the research findings, while Figure 8 provides a revised version of the research model. Table 60 is intended to show which variables are "high-leverage" – i.e., appear to impact multiple outcome measures – versus those that have more isolated effects – i.e., are only related to one outcome measure. In Table 60, the overall direction of the relationship between predictors and outcomes is shown, with a "+" denoting a positive relationships and a "-" denoting a negative relationship. However, Table 60 does not differentiate between direct versus indirect relationships (see Chapters 4 and 5).

Table 60. Summary of Relations Found in this Research

	Attitude	Task KSA	Impact on Area	Overall Perceived Success	% of Goals Met (continuous & dichotomous)
Management Support	+		+	+	
Goal Difficulty		+	-		-
Team Autonomy		+	+		
Goal Clarity	+	+			
Internal Processes	+	+			
Work Area Routineness		+	+		
Team Kaizen Experience		-			-
Team Leader Experience		-			-
Action Orientation			+		-
Functional Heterogeneity	+				
Affective Commitment to Change		+			
Tool Quality				+	
Event Planning Process					+
Tool Appropriateness					

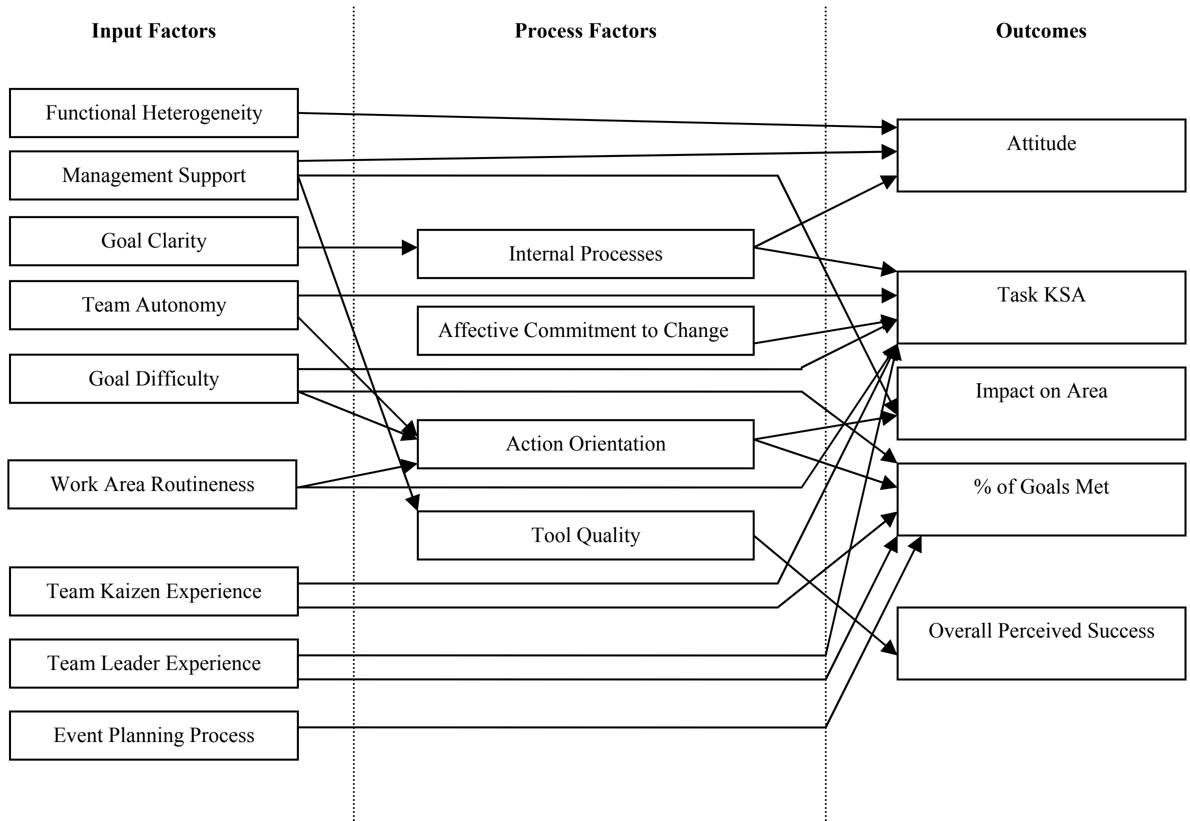

Figure 8. Revised Research Model

As shown, *Management Support* and *Goal Difficulty* have the largest number of common relationships, with significant relationships to three out of four outcome variables. In addition, several other predictors are significantly related to two outcome variables. Only four significant predictors (*Team Functional Heterogeneity*, *Affective Commitment to Change*, *Tool Quality* and *Event Planning Process*) are only significantly related to one outcome variable.

Although the direction of causality in this research is inferred based on theory and methods, rather than definitively established by experimental control, in general, the findings suggest the following guidelines for organizational personnel who plan and/or facilitate Kaizen events:

- First, particular attention should be paid to the *event planning* stages. This research suggests that organizations should not short-change this step in order to hold events more quickly and/or to conduct more events. This research not only demonstrated a direct relationship between the amount of time spent planning (*Event Planning Process*) and *% of Goals Met*, the research also suggests that several other variables that are largely determined by the event planning process have a significant impact on outcomes –

e.g., *Goal Clarity*, *Team Autonomy*, *Management Support*, *Team Functional Heterogeneity*, *Team Kaizen Experience*, and *Team Leader Experience.*

- Second, organizations should recognize the key characteristic of *proposed candidate events* that are significantly related to outcomes – e.g., the difficulty of the specified event objectives (*Goal Difficulty*) and the complexity of the target work area (*Work Are Routineness*).
- Finally, during the event process, particular attention should be given to establishing favorable *Internal Processes*, an appropriate level of *Action Orientation* and high tool application quality (*Tool Quality*).

Several key variables are largely determined – or at least highly influenced – during the *event planning* stage, which occurs after the candidate event has been selected by management:

- For instance, *Goal Clarity* had a significant positive relationship to both *Attitude* and *Task KSA*, indirectly through *Internal Processes*. This suggests that care must be taken to ensure that the goals of the event have been carefully scoped such they are clear to all participating members. This further suggests that perhaps organizations may benefit from additional time devoted to clarifying event objectives and perhaps additional tools – such as "sensing sessions" or a formal event charter. This finding also suggests that recommendations in the Kaizen event literature about allowing team members to develop or substantially refine the event goals during the event may not be advisable.
- *Team Autonomy*, which can be determined in advance of the event but must also be maintained during the event through the facilitation process, had a positive relationship to *Task KSA* and *Impact on Area*, suggesting that allowing teams more freedom in the solution process has both positive social system (human resource) benefits and positive technical system benefits.
- In addition, *Management Support* was significantly positively related to *Attitude*, *Impact on Area* and *Overall Perceived Success*. Several aspects of this variable can be established through the event planning process – e.g., using a standard Kaizen equipment cart, negotiating availability of "offline" support resources during the event, and communication with top management to ensure buy-in to the event and to plan the level of management interaction with the team.
- The final findings of this research related to event planning relate to team composition (i.e., team selection), which also occurs as part of the planning process. *Team Functional Heterogeneity* had a significant negative relationship to *Attitude* and this tradeoff should be considered in event design to allow

countermeasures to be employed to counteract the potential negative effects. *Team Kaizen Experience* and *Team Leader Experience* had negative relationships to *Task KSA* and again this apparent tradeoff in social system outcomes should be considered and countermeasures should be taken. It is noted that *Team Kaizen Experience* and *Team Leader Experience* also had a negative relationship to *% of Goals Met* and, although the nature of this counterintuitive relationship is not yet understood, organizations may want to consider that increasing *Team Kaizen Experience* or *Team Leader Experience* may not necessarily increase goal achievement.

This research found that two key characteristics of proposed candidate events are significantly related to outcomes. These characteristics are more strategic in nature and may not be as directly controllable by the facilitators as other event characteristics, because management often makes the final decisions about what events are held:

- *Goal Difficulty* had a significant positive relationship to *Task KSA* but a significant negative relationship to goal achievement (*% of Goals Met*) and *Impact on Area*. Thus, it appears that organizations face a trade-off in the social system (human resources) benefits achieved from holding more difficult events and the likely level of technical system outcomes. However, countermeasures can be employed to reduce the negative effects of *Goal Difficulty* on technical system outcomes (see the discussion in Chapter 5).
- *Work Area Routineness* had a significant positive relationship to both *Task KSA* (direct) and *Impact on Area* (indirect through *Action Orientation*). Thus, less complex work areas appear to provide favorable "learning laboratories" in that it appears that team members are better able to directly apply lean tools and concepts and see immediate results. It should be noted that difficult ("stretch") goals could be established even in routine work areas; therefore there is not an interchangeable relationship between *Work Area Routineness* and *Goal Difficulty*. However, organizational personnel should balance the benefits of selecting less complex work areas with the strategic impact of events. It is possible, even likely, that more complex work areas – particularly those in knowledge processes, such as engineering – may be the overall bottlenecks in organizational improvement (Murman et al., 2002). This suggests it may be advisable to continue to hold events that target these work areas, despite the tradeoffs with *Task KSA* and *Impact on Area*, but to employ countermeasures to compensate for work area complexity. Or, these work areas could be addressed using another change mechanism – i.e., Six Sigma or TQM/CPI teams.

Finally, several key characteristics of the event process – i.e., activities occurring during the event – were found to be significantly related to event outcomes:

- *Internal Processes*, i.e., harmonious team interactions – particularly consisting of open communication and respect for persons – appears to be a key variable in achieving favorable social system outcomes –i.e., increased *Attitude* and *Task KSA*. *Internal Processes* had a direct relationship to both *Attitude* and *Task KSA* and also appears to mediate the impact of *Goal Clarity* on the outcomes. As discussed in Chapter 5, there are several steps facilitators and team leaders can take in enabling *Internal Processes*, such as establishing ground rules, selecting team members who are "good team players," and assisting in keeping team discussions on track. In addition, as mentioned, *Goal Clarity* appears to act through *Internal Processes* and can therefore be used to increase the level of *Internal Processes*.
- *Action Orientation* is another variable involved in an apparent tradeoff between outcome measures. *Action Orientation* was significantly associated with increased *Impact on Area* but decreased *% of Goals Met*. These findings suggest that facilitators and teams may face a delicate balance between allowing enough "offline" time for training, discussing the problem, brainstorming and otherwise analyzing the problem, and planning solution implementation, and allowing enough time in the target work area for direct observation and implementation (in cases where implementation is possible). This suggests that design guidelines in the Kaizen event practitioner literature that advise the team should spend as much time as possible in the target work do not appear to be a suitable universal recommendation. Although spending the majority of event time in the target work area does increase perceptions of *Impact on Area*, it does not appear to ensure overall event effectiveness in terms of level of goal achievement (*% of Goals Met*).
- Finally, the overall quality of the team's use of problem-solving tools (*Tool Quality*) directly predicted facilitator perceptions of overall success. Thus, it appears that the quality of the team's tool use should also be carefully managed during the event. This could likely be impacted by the quality of the initial training (not measured in this study) and well as directly through the facilitation process. In addition, this research suggests that *Management Support* acts to increase *Tool Quality* and to indirectly influence *Overall Perceived Success* through *Tool Quality*.

It is also interesting to note here that *Tool Appropriateness* was the only predictor variable that was not found to have a significant direct or indirect effect to any of the outcome measures. Conceptually, this would not necessarily seem to imply that *Tool Appropriateness* is not important to event success. Instead, it seems likely that *Tool Appropriateness* is not significant due to relatively low range of variation in the data set. Specifically, out of the 51 events in the data set, only two had a *Tool Appropriateness* rating less than 5.0 and these two events both had ratings greater than 4.5. Upon closer consideration, this relative lack of variation makes sense because the event facilitators, who also provided the *Tool Appropriateness* ratings, often have at least some role in selecting the tools to be used in a particular event, either by directly selecting the tools or by assisting the team in the tool selection process. Thus, it appears unlikely that there would be many events for which the facilitator would rate the average *Tool Appropriateness* as particularly low, because the facilitator would likely, although not necessarily, intervene if he or she perceived the team's tool selection as inappropriate. This perhaps could be addressed in future research by collecting a *Tool Appropriateness* rating from a second data source – e.g., perhaps another facilitator in the organization or a manager who is experienced in continuous improvement activities. However, this could be difficult in some cases where there is only one Kaizen event facilitator and would also add to the burdens of both organizations and researchers in collecting this data. In addition, it is possible that the researchers or other external experts could make judgments about *Tool Appropriateness* based on the report out file and Event Information Sheet results; however, it is not clear that enough information would be provided from these documents to provide solid groundwork for a consistent rating across organizations, particularly since the report out formats – and associated level of detail – differ across organizations. Finally, the impact of *Tool Appropriateness* could perhaps be investigated in post-hoc analysis using Data Envelopment Analysis (DEA) (Charnes et al., 1978) to determine whether events with the lowest efficiency scores demonstrate any patterns on *Tool Appropriateness* – e.g., lower than average scores or at least one tool with a relatively low appropriateness rating, etc.

6.2 Additional Testing of Model Robustness

Although the conclusions in this research were based on a relatively large sample of events at the team-level (i.e., 51 events), the sample at the organizational level was relatively small (i.e., six organizations). In addition, the organizations are all manufacturing organizations of some type, so no purely service organizations were tested, although the sample did include several events in non-manufacturing processes. Additional testing of the robustness of model results can be achieved by sampling additional organizations, including some organizations that only or primarily conduct knowledge work. This will also increase the relative sample size of non-manufacturing versus manufacturing events.

The primary output of this research has been the identification of the most significant set of predictors for each outcome measure. However, additional research is needed to further test and further understand the nature of these apparent relationships. As mentioned in Chapter 5, the potential of conducting field quasi-experiments based on this research appears promising. A baseline of organizational performance could be taken before introducing one of the design suggestions indicated in this research as having positive effects and the impact of changes to the organization's Kaizen process could then be measured.

6.3 Testing of Additional Model Parameters

There were several potential input and process variables that were not tested in the research (see Chapter 2, especially Table 1). Data were collected on some of these variables – e.g.,. actual problem-solving tools used, training on lean tools and techniques, diversity in *Team Kaizen Experience*, etc. – that will enable some additional post-hoc analyses. However, additional variables of interest could also be introduced into the revised research model for testing in future research. For instance, even though some data were collected on Kaizen event training duration and content, future research could explicitly measure team member perceptions of Kaizen event training adequacy. This could be introduced as an additional question in the *Management Support* construct – if the primary interest is in the adequacy of the general support provided to the Kaizen event team – or it could be measured as a stand-alone construct. However, care must be taken to distinguish questions related to the adequacy of the training session during the Kaizen event (content, clarity, etc.) from team member perceptions of the impact of the event on their KSAs. Thus, it would be preferable to measure perceptions of training adequacy prior to measuring event outcomes and to make sure the questions only reflect overall training adequacy and clearly refer to the training session at the beginning of the event.

This research revealed several new variables that may be interrelated with the predictor variables studied in this research and/or may be predictors of event outcomes. These variables include direct measures of event scope, event priority (i.e., importance to top management), team effort, team "spirit," enjoyment of working with peers, and comfort in working with people from different functions. In addition, it may be desirable to collect facilitator, as well as team member perceptions of *Goal Difficulty*. Furthermore, it is suggested that *Overall Perceived Success* be expanded into a multi-item measure in future research.

Finally, this research did not include any explanatory variables that were explicitly measured at the organizational-level. This was due to the fact that this research focused on identifying what team-level explanatory variables predict event outcomes. However in future research, data could be collected on additional organizational-level variables expected to contribute to positive event outcomes – e.g., organizational commitment to Kaizen events, organizational trust, etc. In order to build a hierarchical model, a substantially larger sample size at the organizational level than the original six organizations studied is needed (see the discussion in Chapter 4 and Chapter 5).

6.4 Research on Sustainability of Event Outcomes

This research focused on measuring the initial outcomes of Kaizen events and identifying what event input and event process factors are most strongly related to outcomes. This research did not aim to measure the sustainability of event outcomes or identify factors that support or inhibit sustainability. For organizations to make effective use of Kaizen events as an improvement vehicle, two things are needed. First, organizations must be able to effectively execute events in order to consistently generate positive outcomes – both technical system outcomes and social system outcomes. The outcomes of this research – i.e. the proposed guidelines for organizations – can help organizations with this objective. Second, organizations must be able to largely sustain outcomes in the long-term. This research did not address this objective, and additional research is needed to measure the sustainability of event outcomes in organizations and to identify factors related to sustainability. As has been mentioned, there is an ongoing research initiative at OSU and VT to study the sustainability of event outcomes. This initiative and the research described in this document are both part of a three year study which received funding from NSF in 2005. The first study year (2005-2006) focused on initial outcomes and contained the work described in this document. The next two study years (2006-2007 and 2007-2008) will focus on the sustainability of event outcomes.

REFERENCES

Adam, EE., Jr. (1991), "Quality Circle Performance," *Journal of Management*, Vol. 17 No. 1, pp. 25-39.

Adams, M., Schroer, B.J., & Stewart, S.K. (1997), "Quickstep^{TM} Process Improvement: Time-Compression as a Management Strategy," *Engineering Management Journal*, Vol. 9 No. 2, pp. 21-32.

Alwin, D.F. & Hauser, R.M. (1975), "The Decomposition of Effects in Path Analysis," *American Sociological Review*, Vol. 40, pp. 37-47.

Amabile, T.M. (1989), "The Creative Environment Scales: Work Environment Inventory," *Creativity Research Journal*, Vol. 2, pp. 231-253.

Amason, A.C. & Schweiger, D.M. (1994), "Resolving the Paradox of Conflict, Strategic Decision Making and Organizational Performance," *International Journal of Conflict Management*, Vol. 5, pp. 239-253.

Ancona, D. & Caldwell, D.F. (1992), "Demography and Design: Predictors of New Product Team Performance," *Organization Science*, Vol. 3, pp. 321-341.

Anderson, N. & West, M. (1996), "The Team Climate Inventory: The Development of the TCI and its Application in Team Building for Innovativeness," *European Journal of Work and Organizational Psychology*, Vol. 5, pp. 53-66.

Anderson, N. & West, M. (1998), "Measuring Climate for Work Group Innovation: Development and Validation of the Team Climate Inventory," *Journal of Organizational Behavior*, Vol. 19, pp. 235-258.

Atkinson, J. (1958), "Toward Experimental Analysis of Human Motivation in Terms of Motives, Expectancies and Incentives." *Motives in Fantasy, Action and Society,* Atkinson, J., ed., Princeton, NJ: Van Nostrand, pp. 288-305.

Azzi, A.E. (1993), "Implicit and Category-Based Allocations of Decision-Making Power in Majority-Minority Relations," *Journal of Experimental Psychology*, Vol. 29, pp. 203-228.

Baker, P. (2005), "The Value of Learning" *Works Management*, Vol. 58 No. 3, pp. 19.

Bane, R. (2002), "Leading Edge Quality Approaches in Non-Manufacturing Organizations," *Quality Congress. ASQ's … Annual Quality Proceedings*, pp. 245-249

Baron, R.A. (1990), "Countering the Effects of Destructive Criticism: The Relative Efficacy of Four Interventions," *Journal of Applied Psychology*, Vol. 75, pp. 235-245.

Baron, R.M. & Kenny, D.A. (1986), "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic and Statistical Considerations," *Journal of Personality and Social Psychology*, Vol. 51, pp. 1173-1182.

Bartko, J.J. (1976), "On Various Intraclass Correlation Reliability Coefficients," *Psychological Bulletin*, Vol. 83 No. 5, pp. 762-765.

Bassiri, D. (1998), "Large and Small Sample Properties of the Maximum Likelihood Estimates for the Hierarchical Linear Model," Unpublished Doctoral Dissertation, Michigan State University: East Lansing, MI

Bateman, N. (2005), "Sustainability: The Elusive Element in Process Improvement," *International Journal of Operations and Production Management*, Vol. 25 No. 3, pp. 261-276.

Bateman, N. & David, A. (2002), "Process Improvement Programmes: A Model for Assessing Sustainability," *International Journal of Operations and Production Management*, Vol. 22 No. 5, pp. 515-526.

Batt, R. & Appelbaum, E. (1995), "Worker Participation in Diverse Settings: Does the Form Affect the Outcome, and If So, Who Benefits?" *British Journal of Industrial Relations*, Vol. 33 No. 3, pp. 353-378.

Beckett, A. J., Wainwright, C. E. R., & Bance D. (2000), "Implementing an Industrial Continuous Improvement System: A Knowledge Management Case Study," *Industrial Management & Data Systems*, Vol. 100 No. 7, pp. 330-338.

Belassi, W. & Tukel, O.I. (1996), "A New Framework for Determining Critical Success/Failure Factors in Projects," *International Journal of Project Management*, Vol. 14 No. 3, pp. 141-151.

Bicheno, J. (2001), "Kaizen and Kaikaku." *Manufacturing Operations and Supply Chain Management: The LEAN Approach,* Taylor, D, and Brunt, D., eds., London, UK: Thomson Learning, pp. 175-184.

Bititci, U.S., Turner, T., & Begemann, C. (2000), "Dynamics of Performance Measurement Systems," *International Journal of Project Management*, Vol. 20 No. 6, pp. 692-704.

Blalock, H.M., Jr. (1972), *Social Statistics*, 2nd Edition, New York: McGraw Hill Companies, Inc.

Bliese, P.D. (2000), "Within-Group Agreement, Non-Independence, and Reliability: Implications for Data Aggregation and Analysis." *Multilevel Theory, Research and Methods in Organizations,* Klein, K. J., and Kozlowski, S.W.J., eds., San Francisco: Jossey-Bass, pp. 349-381.

Bliese, P.D. & Halverson, R.R. (1998), "Group Size and Measures of Group-Level Properties: An Examination of Eta-Squared and ICC Values," *Journal of Management*, Vol. 24 No. 2, pp. 157-172.

Bliese, P.D., Halverson, R.R., & Rothberg, J.M. (1994), "Within-Group Agreement Scores: Using Resampling to Estimate Expected Variance," *Academy of Management Best Paper Proceedings*, 303-307.

Bliese, P.D. & Hanges, P.J. (2004), "Being Too Liberal and Too Conservative: The Perils of Treating Grouped Data as Though They Were Independent," *Organizational Research Methods*, Vol. 7 No. 4, pp. 400-417.

Bodek, N. (2002), "Quick and Easy Kaizen," *IIE Solutions*, Vol. 34 No. 7, pp. 43-45.

Bowen, D.E. & Youngdahl, W.E. (1998), "'Lean' Service: In Defense of a Production-Line Approach," *International Journal of Service Industry Management*, Vol. 9 No. 3, pp. 207-225.

Bozzone, V. (2002). *Speed to Market: Lean Manufacturing for Job Shops*, 2nd ed., Amacom: New York.

Bradley, J.R. & Willett, J. (2004), "Cornell Students Participate in Lord Corporation's Kaizen Projects," *Interfaces*, Vol. 34 No. 6, pp. 451-459.

Breitkopf, C.R. (2006), "Perceived Consequences of Communicating Organ Donor Wishes: An Analysis of Beliefs about Defending One's Decision," *Psychology and Health*, Vol. 21 No. 4, pp. 481- 497.

Breslow, N. (1990), "Tests of Hypotheses in Overdispersed Poisson Regression and Other Quasi-Likelihood Models," *Journal of the American Statistical Association*, Vol. 85 No. 410, pp. 565- 571.

Brunet, A. P. & New, S. (2003), "Kaizen in Japan: An Empirical Study," *International Journal of Operations and Production Management*, Vol. 23 No.12, pp. 1426-1446.

Burke, M.J. & Dunlap, M.S. (2002), "Estimating Interrater Agreement with the Average Deviation Index: A User's Guide," *Organizational Research Methods*, Vol. 5 No.2, pp. 159-172.

Burke, M.J., Finkelstein, L.M., & Dusig, M.S. (1999), "On Average Deviation Indices for Estimating Interrater Agreement," *Organizational Research Methods*, Vol. 2 No.1, pp. 49-68.

Butterworth, C. (2001), "From Value Stream Mapping to Shop Floor Improvement: A Case Study of Kaikaku." *Manufacturing Operations and Supply Chain Management: The LEAN Approach,* Taylor, D, and Brunt, D., eds., London, UK: Thomson Learning, pp. 185-193.

Campion, M.A., Medsker, G.J., & Higgs, A.C. (1993), "Relations between Work Group Characteristics and Effectiveness: Implications for Designing Effective Work Groups," *Personnel Psychology*, Vol. 46, pp. 823-850.

Chan, D. (1998), "Functional Relationships Among Constructs in the Same Content Domain at Different Levels of Analysis: A Typology of Composition Models," *Journal of Applied Psychology,* Vol. 83 No. 2, pp. 234-246.

Chang, Y. (2000), "Residual Analysis of the Generalized Linear Models for Longitudinal Data," *Statistics in Medicine,* Vol. 19 No. 10, pp. 1277-1293.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978), "Measuring the Efficiency of Decision Making Units," *European Journal of Operational Research,* Vol. 2, pp. 429-444.

Chow-Chua, C. & Goh, M. (2000), "Quality Improvement in the Healthcare Industry: Some Evidence from Singapore," *International Journal of Health Care Quality Assurance*, Vol. 13 No. 5, pp. 223-229.

Clark, K. (2004), "The Art of Counting," *Chain Store Age*, Vol. 80 No. 10, pp. 94-96 (2 pgs).

Cohen, S.G. & Bailey, D.E. (1997), "What Makes Team Work: Group Effectiveness Research from the Shop Floor to the Executive Suite," *Journal of Management*, Vol. 23 No. 3, pp. 239-290.

Cohen, S.G. & Ledford, G.E. (1994), "The Effectiveness of Self-Managing Teams: A Quasi-Experiment," *Human Relations*, Vol. 47 No. 1, pp. 13-43.

Cohen, S.G., Ledford, G.E., & Spreitzer, G.M. (1996), "A Predictive Model of Self-Managing Work Team Effectiveness," *Human Relations*, Vol. 49 No. 5, pp. 643-676.

Colonge, J.B., Carter, R.L., Fujita, S., & Ban, S. (1993), "Application of Generalized Estimating Equations to a Study of In Vitro Radiation Sensitivity," *Biometrics*, Vol. 49 No. 3, pp. 927-934.

Cooney, R. (2002), "Is 'Lean' a Universal Production System? Batch Production in the Automotive Industry," *International Journal of Operations & Production Management*, Vol. 22 No. 10, pp. 1130-1147.

Cordery, J.L., Mueller, W.S., & Smith, L.M. (1991), "Attitudinal and Behavioral Effects of Autonomous Group Working: A Longitudinal Field Setting," *Academy of Management Journal*, Vol. 34 No. 2, pp. 464-476.

Creswell, J. (2001), "America's Elite Factories," *Fortune*, Vol. 144, No. 4, pp. 206A (6 pgs).

Cronbach, L. (1951), "Coefficient Alpha and the Internal Structure of Tests," *Psychiatrika*, Vol. 16, pp. 297-334.

Cronbach, L. & Meehl, P. (1955), "Construct Validity in Psychological Tests," *Psychological Bulletin*, pp. 281-302.

Crossan, M.M., Lane, H.W., & White, R.E. (1999), "An Organizational Learning Framework: From Intuition to Institution," *Academy of Management Review*, Vol. 24 No. 3, pp. 522-537.

Cuscela, K. N. (1998), "Kaizen Blitz Attacks Work Processes at Dana Corp.," *IIE Solutions*, Vol. 30 No. 4, pp. 29-31.

Dansereau, F., Alutto, J.A., & Yammarino, F.J. (1984), *Theory Testing in Organizational Behavior: The Varient Approach*, Englewood Cliffs, NJ: Prentice-Hall.

David, I. (2000), "Drilled in Kaizen," *Professional Engineering*, Vol. 13 No. 9, pp. 30-31.

Davison, M.L., Kwak, N., Seo, Y.S., & Choi, J. (2002), "Using Hierarchical Linear Models to Examine Moderator Effects: Person-by-Organization Interactions," *Organizational Research Methods*, Vol. 5 No. 3, pp. 231-254.

DeCoster, J. (2002), "Using ANOVA to Examine Data from Groups and Dyads," Available URL: http://www.stat-help.com/notes.html, October 17, 2006.

DeGreene, K. (1973), *Sociotechnical Systems*, Englewood Cliffs, NJ: Prentice-Hall.

Demers, J. (2002), "The Lean Philosophy," *CMA Management*, Vol. 76 No. 7, pp. 31-33.

DeVellis, R. F. (1991), *Scale Development: Theory and Application*, Newbury Park, CA: Sage Publishing.

Doolen, T.L., Hacker, M.E., & Van Aken, E.M. (2003a), "The Impact of Organizational Context on Work Team Effectiveness: A Study of Production Teams," *IEEE Transactions on Engineering Management*, Vol. 50 No. 3, pp. 285-296.

Doolen, T.L., Worley, J., Van Aken, E. M., and Farris, J. (2003b), "Development of an Assessment Approach for Kaizen Events," *Proceedings of the 2003 Industrial Engineering and Research Conference*, Portland, OR, May 18-20, 2003, CD-ROM.

Drickhamer, D. (2004a), "Just-In-Time Training," *Industry Week*, Vol. 253 No. 7, pp. 69.

Drickhamer, D. (2004b), "Braced for the Future," *Industry Week*, Vol. 253 No. 10, pp. 51-52.

Drum, M. & McCullagh, P. (1993), "[Regression Models for Discrete Longitudinal Responses]: Comment," *Statistical Science*, Vol. 8 No. 3, pp. 300-301.

Druskat, V.U. & Pescosolido, A.T. (2000), "The Context of Effective Teamwork Mental Models in Self-Managing Teams: Ownership, Learning, and Heedful Interrelating," *Proceedings of the 2000 Academy of Management Annual Meeting*, Toronto, August 4-9, 2000.

Duncan, G.M. (1972), "Characteristics of Organizational Environments and Perceived Environmental Uncertainty," *Administrative Science Quarterly*, Vol. 17, pp. 313-327.

Dunlap, M.S., Burke, M.J., & Smith-Crowe, K. (2003), "Accurate Tests of Statistical Significance for r_{wg} and Average Deviation Interrater Agreement Indexes," *Journal of Applied Psychology*, Vol. 88 No.2, pp. 356-362.

Durham, C.C., Knight, D., & Locke, E.A. (1997), "Effects of Leader Role, Team-Set Goal Difficulty, Efficacy, and Tactics on Team Effectiveness," *Organizational Behavior and Human Decision Processes*, Vol. 72 No. 2, pp. 203-231.

Earley, P.C., Connolley, T., & Ekegren, G. (1989), "Goals, Strategy Development and Task Performance: Some Limits on the Efficacy of Goal Setting," *Journal of Applied Psychology*, Vol. 74 No.1, pp. 24-33.

Earley, P.C. & Mosakowski, E. (2000), "Creating Hybrid Team Cultures: An Empirical Test of Transnational Team Functioning," *Academy of Management Journal*, Vol. 43 No.1, pp. 26-49.

Easton, G. & Jarrell, J.D. (1998), "The Effects of Total Quality Management on Corporate Performance: An Empirical Investigation," *Journal of Business*, Vol. 71 No. 2, pp. 253-307.

Edmondson, A.C. (2002), "The Local and Variegated Nature of Learning in Organizations: A Group-Level Perspective," *Organization Science*, Vol. 13, No. 2.

Emery, F.E. & Trist, E.L. (1960), "Sociotechnical Systems." *Management Sciences: Models and Techniques,* Churchman, C.W., and Verhulst, M., eds., Oxford, UK: Pergamon, pp. 83-97.

Emrich, L.J. & Piedmonte, M.R. (1992), “On the Small Sample Properties of the Generalized Estimating Equation Estimates for Multivariate Dichotomous Outcomes,” *Journal of Statistical Computing and Simulation*, Vol. 41, pp. 19-29.

Erez, M. & Zidon, I. (1984), “Effects of Goal Acceptance on the Relationship of Goal Setting and Task Performance,” *Journal of Applied Psychology*, Vol. 69, pp. 69-78.

Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., & Strahan, E.J. (1999), “Evaluating the Use of Exploratory Factor Analysis in Psychological Research,” *Psychological Methods*, Vol. 4 No. 3, pp. 272-299.

Farris, J. & Van Aken, E.M. (2005), “Benchmarking Kaizen Event Success: Best Practices Report,” Technical Report 05D-02 (41 pages). Blacksburg, Virginia: Virginia Polytechnic Institute and State University, Department of Industrial and Systems Engineering.

Farris, J., Van Aken, E.M., Doolen, T. L., & Worley, J. (2004), “Longitudinal Analysis of Kaizen Event Effectiveness,” *Proceedings of the 2004 Industrial Engineering and Research Conference*, Houston, TX, May 15-19, 2004, CD-ROM.

Farris, J., Van Aken, E. M., Doolen, T. L., and Worley, J. (2006), “Learning from Kaizen Events: A Research Methodology for Determining the Characteristics of More – and Less – Successful Events.” *Proceedings of the 2006 American Society for Engineering Management Conference*, Huntsville, AL, October 25-28, 2006, CD-ROM.

Field, A. (2005), *Discovering Statistics Using SPSS.* 2nd edition. London: Sage Publications Ltd.

Finch, H. (2006), “Comparison of the Performance of Varimax and Promax Rotations: Factor Structure Recovery for Dichotomous Items,” *Journal of Educational Measurement*, Vol. 43 No. 1, pp. 39-52.

Foley, S., Linnehan, F., Greenhaus, J.H., & Weer, C.H. (2006), “The Impact of Gender Similarity, Racial Similarity, and Work Culture on Family-Supportive Supervision,” *Group & Organization Management*, Vol. 31 No. 4, pp. 420-441.

Foreman, C.R. & Vargas, D.H. (1999), “Affecting the Value Chain through Supplier Kaizen,” *Hospital Materiel Management Quarterly*, Vol. 20 No. 3, pp. 21-27.

“Get Smart, Get Lean” (2003), *Upholstery Design & Management*, Vol. 16 No. 8, pp. 15-19.

George, J.M. (1990), “Personality, Affect, and Behavior in Groups,” *Journal of Applied Psychology*, Vol. 75, pp. 107–116.

George, J.M. & Zhou, J. (2001), “When Openness to Experience and Conscientiousness are Related to Creative Behavior,” *Journal of Applied Psychology*, Vol. 75, pp. 107–116.

Gibson, C. & Vermeulen, F. (2003), “A Healthy Divide: Subgroups as a Stimulus for Team Learning Behavior,” *Administrative Science Quarterly*, Vol. 48 No. 2, pp. 202–239.

Gregory, A. (2003), “Haring for Change,” *Works Management*, Vol. 56 No. 5, pp. 18-21 (3 pgs).

Griffin, R.W. (1988), “Consequences of Quality Circles in an Industrial Setting: A Longitudinal Assessment,” *Academy of Management Journal*, Vol. 31 No. 2, pp. 338-358.

Groesbeck, R. (2001), “An Empirical Study of Group Stewardship and Learning: Implications for Work Group Effectiveness,” Unpublished Doctoral Dissertation, Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg: VA.

Gulbro, R. D., Shonesy, L. & Dreyfus, P. (2000) “Are Small Manufacturers Failing the Quality Test?” *Industrial Management & Data Systems*, Vol. 100 No. 2, pp. 76-80.

Gunsolley, J.C., Getchell, C., & Chinchilli, V.M.. (1995), "Small Sample Characteristics of Generalized Estimating Equations," *Communications in Statistics -- Simulation*, Vol. 24, pp. 869–878.

Gupta, P.P., Dirsmith, M.W., & Fogarty, T.J. (1994), "Coordination and Control in a Government Agency: Contingency and Institutional Theory Perspectives on GAO Audits," *Administrative Science Quarterly*, Vol. 39, pp. 264-284.

Haggard, E.A. (1958), *Intraclass Correlation and the Analysis of Variance,* New York: The Dryden Press.

Hammer, D.P., Mason, H.L., Chalmers, R.K., Popovich, N.G., & Rupp, M.T. (1996), "Development and Testing of an Instrument to Assess Behavioral Professionalism of Pharmacy Students," *Journal of Pharmaceutical Education*, Vol. 64, pp. 141–151.

Hammersley, G. & Pinnington, A. (1999), "Employee Response to Continuous Improvement Groups," *The TQM Magazine*, Vol. 11 No. 1, pp. 29–34.

Hanley, J.A., Negassa, A., deB. Edwardes, M.D., & Forrester, J.E. (2003), "Statistical Analysis of Correlated Data Using Generalized Estimating Equations: An Orientation," *American Journal of Epidemiology*, Vol. 157 No. 4, pp. 364–375.

Hardin, J. W. & Hilbe, J.M. (2003), *Generalized Estimating Equations*, Boca Raton, FL: Chapman & Hall/CRC.

Harrington, J. H. (1998), "Performance Improvement: The Rise and Fall of Reengineering," *The TQM Magazine*, Vol. 10 No. 2, pp. 69–71.

Hart, S.H., Moncrief, W.C., & Parasuraman, J. (1989), "An Empirical Investigation of Salespeople's Performance, Effort and Selling Method During a Sales Contest," *Journal of the Academy of Marketing Science*, Vol. 17 No. 1, pp. 29–39.

Harvey, J. (2004), "Match the Change Vehicle and Method to the Job," *Quality Progress*, Vol. 37 No. 1, pp. 41-48.

Hasek, G. (2000), "Extraordinary Extrusions," *Industry Week*, Vol. 249 No. 17, pp. 79-80.

Hayes, B.E. (1994), "How to Measure Empowerment," *Quality Progress*, Vol. 27 No. 2, pp. 41–46.

Heard, E. (1997), "Rapid-Fire Improvement with Short-Cycle Kaizen," *Proceedings of the American Production and Inventory Control Society*, Washington, DC, October 26-29, 1997, pp. 519-523.

Heard, E. (1998), "Continuous Improvement: Putting Commitment to Work through Short-Cycle Kaizen," *National Productivity Review*, Vol. 17 No. 3, pp. 75–79.

Herscovitch, L. & Meyer, J. P. (2002), "Commitment to Organizational Change: Extension of a Three-Component Model," *Journal of Applied Psychology*, Vol. 87 No. 3, pp. 474-487.

Horton, N.J. & Lipsitz, S.R. (1999) "Review of Software to Fit Generalized Estimating Equation Regression Models," *The American Statistician*, Vol. 53 No.2, pp. 160-169.

Hox, J.J. (1994), "Hierarchical Regression Models for Interviewer and Respondent Effects," *Sociological Methods and Research,* Vol. 17 No. 1, pp. 300–316.

Hyatt, D.E. & Ruddy, T.M. (1997), "An Examination of the Relationship between Work Group Characteristics and Performance: Once More into the Breech," *Personnel Psychology*, Vol. 50 No. 3, pp. 553-585.

Imai, M. (1986), *Kaizen: The Key to Japan's Competitive Success,* New York: Random House.

Institute of Industrial Engineers, (2006) "Definition of Industrial Engineering," IIE Bylaws, Article B-II, Purposes Section, Available URL: http://www.iienet2.org/Details.aspx?id=283, December 5, 2006.

iSixSigma LLC, (2005) "Kaizen," Available URL: http://www.isixsigma.com/dictionary/Kaizen-42.htm, April 29, 2005.

Ivancevich, J.M. & McMahon, J. T. (1977), "A Study of Task-Goal Attributes, Higher Order Need Strength, and Performance," *Academy of Management Journal*, Vol. 20 No. 4, pp. 552-563.

Jackson, S. (1992), "Team Composition in Organizations." *Group Process and Productivity,* Worchel, S, Wood, W. and Simpson, J., eds., London, UK: Sage, pp. 1-12.

James, L.R. (1982), "Aggregation Bias in Estimates of Perceptual Agreement," *Journal of Applied Psychology,* 69 (2), pp. 219-229.

James, L.R., Demaree, R.G., & Wolf, G. (1984), "Estimating Within-Group Interrater Reliability with and without Response Bias," *Journal of Applied Psychology,* 69, pp. 3-32.

James, L.R., Demaree, R.G., & Wolf, G. (1993), "r_{wg}: An Assessment of Within-Group Interrater Agreement," *Journal of Applied Psychology,* Vol. 78 No. 2, pp. 306-309.

James, L.R. & Williams, L.J. (2000), "The Cross-Level Operator in Regression, ANCOVA, and Contextual Analysis." *Multilevel Theory, Research and Methods in Organizations,* Klein, K. J., and Kozlowski, S.W.J., eds., San Francisco: Jossey-Bass, pp. 382-424.

James-Moore, S.M. & Gibbons, A. (1997) "Is Lean Manufacture Universally Relevant? An Investigative Methodology," *International Journal of Operations & Production Management*, Vol. 17 No. 9, pp. 899-911.

Jehn, K. & Bezrukova, K. (2004), "A Field Study of Group Diversity, Workgroup Context, and Performance," *Journal of Organizational Behavior*, Vol. 25 No. 6, pp. 703-729.

Jehn, K., Northcraft, G., & Neale, M. (1999), "Why Differences Make a Difference: A Field Study of Diversity, Conflict, and Performance in Workgroups," *Administrative Science Quarterly*, Vol. 44 No. 4, pp. 741-763.

Jetten, J., Duck, J., Terry, D.J., & O'Brien, A. (2002), "Being Attuned to Intergroup Differences in Mergers: The Role of Aligned Leaders for Low-Status Groups," *Administrative Science Quarterly*, Vol. 28 No. 9, pp. 1194 -1201.

Jha, S., Michela, J., & Noori, H. (1999), "Within the Canadian Boundaries: A Close Look at Canadian Industries Implementing Continuous Improvement," *The TQM Magazine*, Vol. 11 No. 3, pp. 188–197.

Jha, S., Noori, H., & Michela, J. (1996), "The Dynamics of Continuous Improvement: Aligning Organizational Attributes for Quality and Productivity," *International Journal of Quality Science*, Vol. 1 No. 1, pp. 19–47.

Jina, J., Bhattacharya, A.K., & Walton, A.D. (1997), "Applying Lean Principles for High Product Variety and Low Volumes: Some Issues and Propositions," *Logistics Information Management*, Vol. 10 No. 1, pp. 5-13

Johnson, D.E. (1998), *Applied Multivariate Methods for Data Analysis.* Pacific Grove, CA: Brooks/Cole Publishing Company.

Judd, C.M. & Kenny, D.A. (1981), "Process Analysis: Estimating Mediation in Treatment Evaluations," *Evaluation Review*, Vol. 5, pp. 602-619.

Jusko, J. (2004), "Lonely at the Top," *Industry Week*, Vol. 253 No. 10, pp. 58-59.

Katz, D. & Kahn, R.L. (1966), *The Social Psychology of Organizations*, New York: Wiley.

Keating, E.K., Oliva, R., Repenning, N.P., Rockart, S. & Sterman, J.D. (1999), "Overcoming the Improvement Paradox," *European Management Journal*, Vol. 17 No. 2, pp. 120-134.

Kenny, D.A. (2006), "Mediation," Available URL: http://davidakenny.net/cm/mediate.htm, November 7, 2006.

Kenny, D.A. & Judd, C.M. (1986), "Consequences of Violating the Independence Assumptions in Analysis of Variance," *Psychological Bulletin*, Vol. 99 No. 3, pp. 422-431.

Kenny, D.A., Kashy, D.A. & Bolger, N. (1998), "Data Analysis in Social Psychology." *The Handbook of Social Psychology,* Gilbert, D., Fiske, S.T., and Lindzey, G., eds., Boston: McGraw Hill, Vol. 1, pp. 233-265.

Kenny, D.A. & La Voie, L. (1985), "Separating Individual and Group Effects," *Journal of Personality and Social Psychology*, Vol. 48 No. 2, pp. 339-348.

"Keys to Success" (1997), *Industry Week*, Vol. 246 No. 16, pp.20-21.

Kirby, K. & Greene B. (2003), "How Value Stream Type Affects the Adoption of Lean Production Tools and Techniques," *Proceedings of the 2003 Industrial Engineering and Research Conference*, Portland, OR, May 18-20, 2003, CD-ROM.

Kirkman, B.L. & Shapiro, D.L. (1997), "The Impact of Cultural Values on Employee Resistance to Teams: Toward a Model of Globalized Self-Managing Work Team Effectiveness," *Academy of Management Review*, Vol. 22, pp. 730-745.

Kirkman, B.L. & Rosen, B. (1999), "Beyond Self-Management: Antecedents and Consequences of Team Empowerment," *Academy of Management Journal*, Vol. 42 No. 1, pp. 58-74.

Klaus, L.A. (1998), "Kaizen Blitz Proves Effective for Dana Corporation," *Quality Progress*, Vol. 31 No. 5, pp. 8.

Klein, K. J. & Kozlowski, S.W.J. (2000), "From Micro to Meso: Critical Steps in Conceptualizing and Conducting Multilevel Research," *Organizational Research Methods*, Vol. 3 No. 3, pp. 211-236.

Kleinsasser, J. (2003), "Kaizen Seminar Brings Change to Library, Purchasing Processes," *Inside WSU*, Available URL: http://webs.wichita.edu/dt/insidewsu/show/article.asp?270, March 20, 2004.

Kline, P. (1994), *An Easy Guide to Factor Analysis*, London: Routledge.

Kosandal, P. & Farris, J. (2004), "The Strategic Role of the Kaizen Event in Driving and Sustaining Organizational Change," *Proceedings of the 2004 American Society for Engineering Management Conference*, Alexandria, VA, October 20-23, 2004, pp. 517-526.

Kotter, J. (1995), "Leading Change: Why Transformation Efforts Fail," *Harvard Business Review,* Vol. 73 No. 2, pp. 59-67.

Kozlowski, S.W.J. & Hattrup, K. (1992), "A Disagreement about Within-Group Agreement: Disentangling issues of consistency versus consensus," *Journal of Applied Psychology*, Vol. 77 No.2, pp. 161-167.

Kozlowski, S.W.J. & Hults, B.M. (1987), "An Exploration of Climates for Technical Updating and Performance," *Personnel Psychology*, Vol. 40, pp. 539-563.

Kumar, S. & Harms, R. (2004), "Improving Business Processes for Increased Operational Efficiency: A Case Study," *Journal of Manufacturing Technology Management,* Vol. 15 No. 7, pp. 662-674.

Laraia, A. (1998), "Change it NOW!" *Manufacturing Engineering*, Vol. 121 No. 4, pp. 152.

Laraia, A. C., Moody, P. E., & Hall, R. W. (1999), *The Kaizen Blitz: Accelerating Breakthroughs in Productivity and Performance*, New York: The Association for Manufacturing Excellence.
Larson, M. (1998a), "Five Days to a Better Process: Are You Ready for Kaizen," *Quality*, Vol. 37 No. 6, pp. 38.

Larson, M. (1998b), "Lantech's Kaizen Diary: Monday through Friday," *Quality*, Vol. 37 No. 6, pp. 40.

Lawal, B. (2003), *Categorical Data Analysis with SAS and SPSS Applications*, London: Lawrence Erlbaum Associates, Inc.

Lawler . E.E. and S.A. Mohrman. (1985), "Quality Circles after the Fad," *Harvard Business Review*, 85, 1, pp. 64-71.

Lawler . E.E. and S.A. Mohrman. (1987), "Quality Circles: After the Honeymoon," *Organizational Dynamics*, 19, pp. 15-26

Lawyer, S.R., Resnick, H.S., Galea, S., Ahern, J., Kilpatrick, D.G., & Vlahov, D. (2006), "Predictors of Peritraumatic Reactions and PTSD Following the September 11th Terrorist Attacks," *Psychiatry* Vol. 69 No. 2, pp. 130-141.

LeBlanc, G. (1999), "Kaizen at Hill-Rom," *Center for Quality of Management Journal*, Vol. 8 No. 2, pp. 49-53.

Leedy, P.D. & Ormrod, J.E. (2005), *Practical Research: Planning and Design*, (8th ed.), Upper Saddle River, NJ: Prentice Hall.

Letens, G., Farris, J. & Van Aken, E. M. (2006), "Development and Application of a Framework for the Design and Assessment of a Kaizen Event Program." *Proceedings of the 2006 American Society for Engineering Management Conference*, Huntsville, AL, October 25-28, 2006, CD-ROM.

Lewis, J.P. (2000), *The Project Manager's Desk Reference*, New York: McGraw Hill Companies, Inc.

Liang, K.Y. & Zeger, S.L. (1986), "Longitudinal Data Analysis using Generalized Linear Models," *Biometrika*, Vol. 73, pp. 13-22.

Lipsitz, S. (1999), Lecture Notes on the GEE for BMTRY 726. Department of Biometry and Epidemiology, MUSC, Charleston, SC.

Locke, E.A. & Latham, G.P. (1990), *A Theory of Goal Setting and Task Performance*, Englewood Cliffs, NJ: Prentice Hall.

Locke, E.A. & Latham, G.P. (2002), "Building a Practically Useful Theory of Goal Setting and Task Motivation: A 35-Year Odyssey," *American Psychologist*, Vol. 57 No. 9, pp. 705-717.

Lovelace, D., Shapiro, D.L., & Weingart, L.R. (2001), "Maximizing Cross-Functional New Product Teams' Innovativeness and Constraint Adherence: A Conflict Communications Perspective," *Academy of Management Journal*, Vol. 44 No. 4, pp. 779-793.

MacKinnon, D.P., Krull, J.L., & Lockwood, C.M. (2000), "Equivalence of the Mediation, Confounding and Suppression Effect," *Prevention Science*, Vol. 1 No. 4, pp. 173-181.

MacKinnon, D.P., West, G., & Dwyer, J.H. (1995), "A Simulation Study of Mediated Effect Measures," *Multivariate Behavioral Research*, Vol. 30 No. 1, pp. 41-62.

Marks, M.L., Mirvis, P.H., Hackett, E.J., & Grady, J.F. (1986), "Employee Participation in a Quality Circle Program: Impact on Quality of Work Life, Productivity, and Absenteeism," *Journal of Applied Psychology*, Vol. 71, pp. 61-69.

Martin, B.A.., Snell, A.F., & Callahan, C.M. (1999), "An Examination of Individual Differences in the Relation of Subjective Goal Difficulty to Performance in a Goal-Setting Model," *Human Performance*, Vol. 12 No.2, pp. 115-135.

Martin, D.W. (1996), *Doing Psychology Experiments*, (4th ed.), Pacific Grove, CA: Brooks/Cole Publishing Company.

Martin, K. (2004), "Kaizen Events: Achieving Dramatic Improvement through Focused Attention and Team-Based Solutions," *Society for Health Systems Newsletter*, August 2004, pp. 6-7.

McNichols, T., Hassinger, R., & Bapst, G.W. (1999), "Quick and Continuous Improvement Through Kaizen Blitz," *Hospital Materiel Management Quarterly*, Vol. 20 No. 4, pp. 1–7.

McGarrie, B. (1998), "Case Study: Production Planning and Control - Selection, Improvement and Implementation," *Logistics Information Management*, Vol. 11 No. 1, pp. 44–52.

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

Melnyk, S. A., Calantone, R. J., Montabon, F. L., & Smith, R. T. (1998), "Short-term Action in Pursuit of Long-Term Improvements: Introducing Kaizen Events," *Production and Inventory Management Journal*, Vol. 39 No. 4, pp. 69-76.

Mika, G. L. (2002), *Kaizen Event Implementation Manual,* 2nd Edition, Wake Forest, NC: Kaizen Sensei.

Miller, B.D. (2004), "The Role of Cognitive Psychology in Lean Manufacturing Research," *Proceedings of the 2004 Industrial Engineering and Research Conference*, Houston, TX, May 15-19, 2004, CD-ROM.

Miltenberg, J. (1995), *Manufacturing Strategy*, Portland, OR: Productivity Press.

Minton, E. (1998), "Profile: Luke Faulstick—'Baron of blitz' has Boundless Vision of Continuous Improvement," *Industrial Management*, Vol. 40 No. 1, pp. 14-21.

Mohr, M. L. & Mohr, H. (1983), *Quality Circles*, Menlo Park, CA: Addison-Wesley Publishing Company.

Molleman, E. (2005), "The Multilevel Nature of Team-Based Work Research," *Team Performance Management*, Vol. 11 No. 3/4, pp. 113-124.

Monden, Y. (1983), *Toyota Production System*, Norcross, GA: Industrial Engineering and Management Press.

Montgomery, D.C., & Runger, G.C. (1999), Applied Statistics and Probability for Engineers (2nd ed.), New York: John Wiley & Sons, Inc.

Muchinsky, P.M. (2000), *Psychology Applied to Work*, (6th ed.), Belmont, CA: Wadsworth/ Thomson Learning.

Murman, E., Allen, T., Bozdogan, K., Cutcher-Gershenfeld, J., McManus, H., Nightingale, E., Rebentisch, E., Shields, T., Stahl, F., Walton, M., Warmkessel, J., Weiss, S., and Widnall, S, (2002), *Lean Enterprise Value*, London: Palgrave.

Neter, J., Kutner, M.H., Nachtsheim, C.J., & Wasserman, W. (1996), *Applied Linear Statistical Models* (4th ed.), New York: McGraw-Hill Companies, Inc.

Nicolini, D. (2002), "In Search of 'Project Chemistry'," *Construction Management and Economics*, Vol. 20 No. 2, pp. 167-177.

Nunnally, J. D. (1978). *Psychometric theory,* (2nd ed.), New York: McGraw-Hill.

Oakeson, M. (1997), “Kaizen Makes Dollars & Sense for Mercedes-Benz in Brazil,” *IIE Solutions*, Vol. 29 No. 4, pp. 32-35.

Pan, W. (1999), “On the Robust Variance Estimator in Generalized Estimating Equations,” *Biometrika*, Vol. 88, pp. 901- 906.

Pan, W. & Wall, M.M. (2002), “Small Sample Adjustments in Using the Sandwich Variance Estimator in Generalized Estimating Equations,” *Statistics in Medicine*, Vol. 21, pp. 1429-1441.

Papadopoulou, T.C. & Ozbayrak, M. (2005), “Leanness: Experiences from the Journey to Date,” *Journal of Manufacturing Technology Management*, Vol. 16 No. 7, pp. 784-807.

Pasmore, W. & King, D. (1978), “Understanding Organizational Change: A Comparative Study of Multifaceted Interventions,” *Journal of Applied Behavioral Sciences*, Vol. 14, pp. 455-468.

Passos, A.M. & Caetano, A. (2003), “Exploring the Effects of Intragroup Conflict and Past Performance Feedback on Team Effectiveness,” *Journal of Managerial Psychology*, Vol. 20 No. 3/4, pp. 231-244.

Patel, S., Dale, B.G., & Shaw, P. (2003), “Set-up Time Reduction and Mistake Proofing Methods: An Examination in Precision Component Manufacturing,” *The TQM Magazine*, Vol. 13 No. 3, pp. 175–179.

Patil, H. (2003), “A Standard Framework for Sustaining Kaizen Events,” Unpublished Master’s Thesis, Department of Industrial and Manufacturing Engineering, Wichita: KS.

Patton, R. (1997), “Two-Day Events: Old Hat at Nissan,” *Industry Week*, Vol. 246 No. 16, pp.24.

Pearson, C.A.L. (1992), “Autonomous Workgroups: An Evaluation at an Industrial Site,” *Human Relations*, Vol. 49 No.5, pp. 905-936.

Pelled, L.H., Eisenhardt, K.M., & Xin, K.R. (1999), “Exploring the Black Box: An Analysis of Work Group Diversity, Conflict and Performance,” *Administrative Science Quarterly*, Vol. 44 No. 1, pp. 1-28.

Perrow, C. (1967), “A Framework for the Comparative Analysis of Organizations,” *American Sociological Review*, Vol. 32, pp. 194-208.

Peugh, J.L. & Enders, C.K. (2005), “Using the SPSS Mixed Procedure to Fit Cross-Sectional and Longitudinal Multilevel Models,” *Educational and Psychological Measurement*, Vol. 65 No. 5, pp. 717-741.

Pickering, J. & Kisangani, E.F. (2005) “Democracy and Diversionary Military Intervention: Reassessing Regime Type and the Diversionary Hypothesis,” *International Studies Quarterly*, Vol. 49, pp. 23-43.

Pinto, J.K. & Mantel, Jr., S.J. (1990), “The Causes of Project Failure,” *IEEE Transactions on Engineering Management*, Vol. 37 No. 4, pp. 269-275.

Pinto, J.K. & Slevin, D.P. (1987), “Critical Factors in Successful Project Implementation,” *IEEE Transactions on Engineering Management*, Vol. 34 No. 1, pp. 22-27.

Pinto, J.K. & Slevin, D.P. (1989), “Critical Success Factors in R&D Projects,” *Research Technology Management*, Vol. 32 No. 1, pp. 31-35.

Poulsen, B.O. (2002), “A Comparison of Bird Richness, Abundance and Trophic Organization in Forests of Ecuador and Demark: Are High-Altitude Andean Forests Temperate or Tropical?,” *Journal of Tropical Ecology*, Vol. 18, pp. 615-636.

Prentice, R.L. (1988), “Correlated Binary Regression with Covariates Specific to Each Binary Observation,” *Biometrics*, Vol. 44 No. 4, pp.1033-1048.

Pritchard, S. (2002), "Brainstorming Teams Boost Productivity," *Works Management*, Vol. 55 No. 4, pp.67.

Project Management Institute, (2000), *A Guide to the Project Management Body of Knowledge (PMBOK® Guide),* 2000 Edition, Newtown Square, PA: Project Management Institute.

Purdum, T. (2004), "A Force to Reckon With," *Industry Week*, Vol. 253 No. 10, pp. 39-40.

Raudenbush, A.S. & Byrk, S.W. (2002), *Hierarchical Linear Models: Applications and Data Analysis Methods*, 2nd Edition, Newbury Park, CA: Sage.

Redding, R. (1996), "Lantech 'Kaizen' Process Draws 63 Observers from Across the Globe," *Business First of Louisville*, Available URL: http://louisville.bizjournals.com/louisville/stories/1996/09/30/story3.html, June 10, 2004.

Reed, K.K., Lubatkin, M., & Srinivasan, N. (2006), "Proposing and Testing an Intellectual Capital-Based View of the Firm," *Journal of Management Studies*, Vol. 42 No. 4, pp. 867-893.

Repenning, N.P. & Sterman, J.D. (2002), "Capability Traps and Self-Confirming Attribution Errors in the Dynamics of Process Improvement," *Administrative Science Quarterly*, Vol. 47, pp. 265-295.

Roby, D. (1995), "Uncommon Sense: Lean Manufacturing Speeds Cycle Time to Improve Low-Volume Production at Hughes," *National Productivity Review*, Vol. 14 No. 2, pp. 79-87.

Rockart, J. F. (1979), "Chief Executives Define Their Own Data Needs," *Harvard Business Review*, Vol. 57 No. 2, pp.81-93.

Roth, P.L. & Switzer, F.S. (1999), "Missing Data: Instrument-Level Heffalumps and Item-Level Woozles," Academy of Management, Research Methods Division *Research Methods Forum*, Vol. 4. Available URL: http://division.aomonline.org/rm/1999_RMD_Forum_Missing_Data.htm, October 18, 2006.

Roth, P.L., Switzer, F.S., III, & Switzer, D.M. (1999), "Missing Data in Multiple Item Scales: A Monte Carlo Analysis of Missing Data Techniques," *Organizational Research Methods*, Vol. 2 No. 3, pp. 211–232.

Rusiniak, S. (1996), "Maximizing Your IE Value," *IIE Solutions*, Vol. 28 No. 6, pp. 12-16.

Sabatini, J. (2000), "Turning Japanese," *Automotive Manufacturing & Production*, Vol. 112 No. 10, pp. 66-69.

Sarin S. & McDermott, C. (2003), "The Effect of Team Leader Characteristics on Learning, Knowledge Application and Performance of Cross-Functional New Product Development Teams," *Decision Sciences*, Vol. 34 No. 4, pp. 707-739.

SAS Institute Inc. (2006), "Generalized Linear Models Theory," *SAS/STAT User's Guide*, SAS OnlineDoc® 9.1.3. Available URL: http://support.sas.com/onlinedoc/913/docMainpage.jsp, November 7, 2006.

Scheaffer, R. L., Mendenhall, W., III., & Ott, R. L. (1996), *Elementary Survey Sampling*, 5th Edition, Belmont, CA: Wadsworth Publishing Company.

Schneider, B., White, S., & Paul, M.C. (1998), "Linking Service Climate and Customer Perceptions of Service Quality: Test of a Causal Model," *Journal of Applied Psychology*, Vol. 83 No.2, pp.150-163.

Schneiderman, A.M. (1988), "Setting Quality Goals," *Quality Progress*, Vol. 21 No. 4, pp.51-57.

Schroeder D.M. & Robinson, A.G. (1991), "America's Most Successful Export to Japan: Continuous Improvement Programs," *Sloan Management Review*, Vol. 32 No. 3, pp. 67-81.

Seers, A., Petty, M.M., & Cashman, J.F. (1995), "Team Member Exchange under Team and Traditional Management," *Group & Organization Management*, Vol. 20 No.1, pp.18-38.

Shannon, C. (1948), "A Mathematical Theory of Communications," *Bell Systems Technical Journal*, Vol. 27, pp. 397-423, 623-656.

Shenhar, A.J., Tishler, A., Dvir, D., Lipovetsky, S., & Lechler, T. (2002), "Refining the Search for Project Success Factors: A Multivariate, Typological Approach," *R&D Management*, Vol. 32 No.2, pp.111-126.

Sheridan, J. H. (1997a), "Guru's View of the Gemba," *Industry Week*, Vol. 246 No. 16, pp. 27-28.

Sheridan, J. H. (1997b), "Kaizen Blitz," *Industry Week*, Vol. 246 No. 16, pp.18-27.

Sheridan, J. H. (2000a), "A New Attitude," *Industry Week*, Vol. 249 No. 10, pp. 16.

Sheridan, J. H. (2000b), "'Lean Sigma' Synergy," *Industry Week*, Vol. 249 No. 17, pp.81-82.

Singer, J.D. (1998), "Using SAS PROC Mixed to Fit Multilevel Models, Hierarchical Models, and Individual Growth Models," *Journal of Educational and Behavioral Statistics*, Vol. 23 No. 4, pp.323-355.

Slevin, D.P. & Pinto, J.K., (1986), "The Project Implementation Profile: New Tool for Project Managers," *Project Management Journal*, Vol. 17 No. 4, pp. 57-70.

Smith, B. (2003) "Lean and Six Sigma – A One-Two Punch," *Quality Progress*, Vol. 36 No. 4, pp. 37-41.

Smith, R.J. (1994) "Degrees of Freedom in Interspecific Allometry: An Adjustment for the Effects of Phylogenetic Constraint," *American Journal of Physical Anthropology*, Vol. 93 No. 1, pp. 95-107.

Sobel, M.E. (1982), "Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models," *Sociological Methodology*, Vol. 13, pp.290-312.

Steel, R.P., Jennings K.R., & Linsey, J.T., (1990), "Quality Circle Problem Solving and Common Cents: Evaluation Study Findings from a United States Mint," *Journal of Applied Behavioral Science*, Vol. 26 No. 3, pp. 365-382.

Sterman, J.D., Repenning, N.P., & Kofman, F. (1997), "Unanticipated Side Effects of Successful Quality Programs: Exploring a Paradox of Organizational Improvement," *Management Science*, Vol. 43, pp. 503-521.

Szinovacz M.E. & Davey, A. (2001), "Retirement Effects of Parent-Adult Child Contacts," *The Gerontologist*, Vol. 41 No. 2, pp. 191-200.

Taninecz, G., (1997), "Cooper Automotive-Wagner Lighting," *Industry Week*, Vol. 246 No. 19, pp. 32 (3 pgs).

Tanner, C. & Roncarti, R. (1994), "Kaizen Leads to Breakthroughs in Responsiveness – and the Shingo Prize – at Critikon," *National Productivity Review*, Vol. 13 No. 4, pp. 517-531.

Taylor, D.L. & Ramsey, R.K. (1993), "Empowering Employees to 'Just Do It'," *Training and Development*, Vol. 47 No. 5, pp. 71-76 (5 pgs).

Teachman, J.D. (1980), "Analysis of Population Diversity," *Sociological Methods and Research*, Vol. 8 No. 3, pp. 341-362.

Tilson, B. (2001), "Success and Sustainability in Automotive Supply Chain Improvement Programmes: A Case Study of Collaboration in the Mayflower Cluster," *International Journal of Innovation Management*, Vol. 5 No. 4, pp. 427-456.

Tinsley, H.E.A. & Tinsley, D.J. (1987), "Uses of Factor Analysis in Counseling Psychology Research," *Journal of Counseling Psychology*, Vol. 34, pp. 414-424.

Treece, J.B. (1993), "Improving the Soul of an Old Machine," *Business Week* Vol. 3342, pp. 134.

Trist, E.L. & Bamforth, K.W. (1951), "Some Social and Psychological Consequences of the Long-Wall Method of Coal-Getting," *Human Relations*, Vol. 4, pp. 3-38.

Trist, E.L., Higgin, G.W., Murray, H., & Pollock, A.B. (1963), *Organizational Choice*, London, UK: Tavistock.

Tukel, O.I. & Rom, W.O. (1998), "Analysis of the Characteristics of Projects in Diverse Industries," *Journal of Operations Management*, Vol. 16 No.1, pp. 43-61.

Van Aken, E.M. & Kleiner, B.M. (1997), "Determinants of Effectiveness for Cross-Functional Organizational Design Teams," *Quality Management Journal*, Vol. 4 No. 2, pp. 51-79.

Van den Bree, M.B.M, Schieken, R.M., Moskowitz, W.B., & Eaves, L.J. (1996), "Genetic Regulation of Hemodynamic Variable During Dynamic Exercise," *Circulation*, Vol. 94, pp. 1864-1869.

Van der Leeden, R. & Busing, F.T.M.A. (1994), "First Iteration versus igls/rgls Estimates in Two-Level Models: A Monte-Carlo Study with ML3," *Psychometrics and Research Methodology*, preprint PRM 94-03.

Van Mierlo, H., Vermunt, J.K., & Rutte, C.G. (2006) "Composing Group-Level Constructs from Individual-Level Survey Data," working paper. Available URL: http:// spitswww.uvt.nl/~vermunt/mierlo2006.pdf, October 4, 2006.

Vasilash, G.S. (1993), "Walking the Talk of Kaizen at Freudenberg-NOK," *Production*, Vol. 105, No. 12, pp. 66-71.

Vasilash, G.S. (1997), "Getting Better—Fast," *Automotive Design & Production*, Vol. 109, No. 8, pp. 66-68.

Vitalo, R. L., Butz, F., & Vitalo, J.P. (2003), *Kaizen Desk Reference Standard*, Hope, ME: Vital Enterprises.

"Waste Reduction Program Slims Fleetwood Down," (2000), *Strategic Direction*, Vol. 16 No. 9, pp. 19-21.

Watson, L. (2002), "Striving for Continuous Improvement with Fewer Resources? Try Kaizen," *Marshall Star*, Nov. 28 2002, pp. 1 (4 pgs).

Weisman, C.S., Gordon, D.L., & Cassard, S.D., (1993), "The Effects of Unit Self-Management on Hospital Nurses' Work Process, Work Satisfaction, and Retention," *Medical Care*, Vol. 31 No. 5, pp. 381-393.

Wheatley, B. (1998), "Innovation in ISO Registration," *CMA*, Vol. 72 No. 5, pp. 23.

Whitford, A.B. & Yates, J. (2003) "Policy Signals and Executive Governance: Presidential Rhetoric in the War on Drugs," *The Journal of Politics*, Vol. 65 No. 4, pp. 995-1012.

Wilson, L., Van Aken, E. M., & Frazier, D. (1998), "Achieving High Performance Work Systems Through Policy Deployment: A Case Application," *Proceedings of the 1998 International Conference on Work Teams*.

Wilson, R. (2005), "Guard the LINE," *Industrial Engineer*, Vol. 37 No. 4, pp. 46-49.

"Winning with Kaizen," (2002), *IIE Solutions*, Vol. 34 No. 4, pp. 10.

Withey, M., Daft, R.L., & Cooper, W.H. (1983) "Measures of Perrow's Work Unit Technology: An Empirical Assessment and a New Scale" *Academy of Management Journal*, Vol. 26 No. 1, pp. 45-63.

Wittenberg, G. (1994), "Kaizen – The Many Ways of Getting Better," *Assembly Automation*, Vol. 14 No. 4, pp. 12+ (6 pgs).

Womack, J. & Jones, D. (1996a), "Beyond Toyota: How to Root out Waste and Pursue Perfection," *Harvard Business Review*, Vol. 74 No. 5, pp. 140-158.

Womack, J. & Jones, D. (1996b), *Lean Thinking: Banish Waste and Create Wealth in Your Corporation*, New York: Simon & Schuster.

Womack, J., Roos, D., & Jones, D. (1990), *The Machine that Changed the World*, New York: Rawson and Associates.

Wood R. & Locke, E.A. (1990), "Goal Setting and Strategy Effects on Complex Tasks." *Research in Organizational Behavior,* Staw, B., and Cummings, L., eds., Greenwich, CT: JAI Press, Vol. 12, pp. 73-109.

Wright, R. (1936), "Factors Affecting the Cost of Airplanes," *Journal of Aeronautical Sciences*, Vol. 3, pp. 122-128.

APPENDIX A: UNCATEGORIZED LIST OF FACTORS FROM KAIZEN EVENT LITERATURE

Factor	Sources
One week or shorter (short duration)	LeBlanc, 1999; Oakeson, 1997; Vasilash, 1997; Drickhamer, 2004b; Watson, 2002; Smith, 2003; Cuscela, 1998; McNichols et al., 1999; Martin, 2004; Sheridan, 1997b; Patton, 1997; Bradley & Willett, 2004; Vasilash, 1993; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998
Linked to organizational strategy	LeBlanc, 1999; "Keys to Success," 1997; Melnyk et al., 1998
Action orientation	LeBlanc, 1999; Redding, 1996; Smith, 2003; Martin, 2004; Sheridan, 1997b; Patton, 1997; Vasilash, 1993; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998
Use cross-functional teams	LeBlanc, 1999; Drickhamer, 2004b; Rusiniak, 1996; Demers, 2002; Smith, 2003; Cuscela, 1998; McNichols et al., 1999; Martin, 2004; Sheridan, 1997b; Vasilash, 1993; Vasilash, 1993; Adams et al., 1997; Melnyk et al., 1998
Teams have implementation authority	LeBlanc, 1999; Oakeson, 1997; Minton, 1998; Martin, 2004; Sheridan, 1997b; Bradley & Willett, 2004; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998
Challenging (stretch) goals	LeBlanc, 1999; Minton, 1998; Rusiniak, 1996; Cuscela, 1998; Bradley & Willett, 2004; Bicheno, 2001
10 – 12 people on Kaizen event team	LeBlanc, 1999; Demers, 2002; Watson, 2002;
Including "fresh eyes" (people with no prior knowledge of the target area) on the team	LeBlanc, 1999; Vasilash, 1997; Kleinsasser, 2003; Minton, 1998; Cuscela, 1998; McNichols et al., 1999; Martin, 2004; Bradley & Willett, 2004; Melnyk et al., 1998
Require a standard, reliable target process/work area as input	LeBlanc, 1999; Bradley & Willett, 2004
Management support essential	Bane, 2002; Hasek, 2000; Vasilash, 1997; Rusiniak, 1996; Cuscela, 1998; Martin, 2004; Sheridan, 1997b; "Keys to Success," 1997; Bradley & Willett, 2004; Vasilash, 1993; Bicheno, 2001; Adams et al., 1997
Spacing out events (e.g., only 1 event per quarter)	Taninecz, 1997
"No layoffs" policy	Redding, 1996; Vasilash, 1997; Creswell, 2001; "Winning with Kaizen," 2002; Womack & Jones, 1996a; "Keys to Success," 1997; Bradley & Willett, 2004; Melnyk et al., 1998
Organization-wide commitment to change	Redding, 1996
Including people from the work area on the Kaizen event team	Redding, 1996; Minton, 1998; Womack & Jones, 1996a; Martin, 2004; Sheridan, 1997b; Bradley & Willett, 2004; Vasilash, 1993; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998
Including outside consultants on the Kaizen event team	Oakeson, 1997; Bicheno, 2001
Including managers and supervisors on the Kaizen event team	Oakeson, 1997; "Keys to Success," 1997; Vasilash, 1993; Bicheno, 2001
Including customers on the Kaizen event team	Hasek, 2000; Vasilash, 1997; McNichols et al., 1999; Vasilash, 1993; Adams et al., 1997; Melnyk et al., 1998
Including suppliers on the Kaizen event team	Vasilash, 1997; McNichols et al., 1999; Vasilash, 1993; Adams et al., 1997; Melnyk et al., 1998
Concurrent Kaizen events	Vasilash, 1997; Watson, 2002; Cuscela, 1998; Bradley & Willett, 2004; Adams et al., 1997
Dedicated room for Kaizen event team meetings	Creswell, 2001
Snacks provided to team during Kaizen event	Creswell, 2001; Adams et al., 1997
Two weeks or shorter	Minton, 1998; Demers, 2002

Focused – on a specific process, product, or problem	Minton, 1998; Drickhamer, 2004b; Martin, 2004; Sheridan, 1997b; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998
Targeted at areas that can provide a "big win" (big impact on organization)	Minton, 1998; Cuscela, 1998; Martin, 2004; Sheridan, 1997b; "Keys to Success," 1997; Bradley & Willett, 2004; Melnyk et al., 1998
Less than two hours of formal training provided to team	Minton, 1998; McNichols et al., 1999
Use of subteams	Minton, 1998; McNichols et al., 1999; Sheridan, 1997b; Bicheno, 2001
Videotapes of SMED	Minton, 1998; Bradley & Willett, 2004
Brainstorming	Minton, 1998; Watson, 2002; Martin, 2004; Bradley & Willett, 2004; Vasilash, 1993
Team members dedicated only to Kaizen event during its duration	Minton, 1998; McNichols et al., 1999; Martin, 2004; Bradley & Willett, 2004; Bicheno, 2001; Melnyk et al., 1998
Including only one employee per department on the Kaizen event team (except for the department being blitzed), to avoid over-burdening any department	Minton, 1998
Facilitators provide "short courses" on topics "on the spot" if a team gets stuck	Minton, 1998
Team members who aren't from the process get training in the process and may even work in the production line for a few days before the Kaizen event	Minton, 1998
Cost is <u>not</u> a factor	Minton, 1998
Repeat Kaizen events in a given work area	"Winning with Kaizen," 2002; Purdum, 2004; Womack & Jones, 1996a; McNichols et al., 1999; Sheridan, 1997b; Bradley & Willett, 2004; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998
Use of a Kaizen newspaper	"Winning with Kaizen," 2002; McNichols et al., 1999; Martin, 2004; Bradley & Willett, 2004; Melnyk et al., 1998
Low cost solutions	Purdum, 2004; Cuscela, 1998; Martin, 2004; Sheridan, 1997b; Vasilash, 1993; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998
Can be held based on employee suggestions for improvement	Jusko, 2004; Watson, 2002
Avoid preconceived solutions	Rusiniak, 1996; Bradley & Willett, 2004
Seek improvement, not optimization	Rusiniak, 1996; Vasilash, 1993
Requires a well-defined problem statement as input	Rusiniak, 1996; Adams et al., 1997
Avoid problems that are too big and/or emotionally involved	Rusiniak, 1996; Sheridan, 1997b
3 – 5 people on Kaizen event team	Rusiniak, 1996
Used to implement lean manufacturing	Vasilash, 1997
Focused on waste elimination	Watson, 2002; Cuscela, 1998; Martin, 2004; Patton, 1997; Adams et al., 1997
Questioning the current process – asking why things are done the way they are	Watson, 2002; Minton, 1998
Team members volunteer to participate	Watson, 2002; Adams et al., 1997
Each team member has specific knowledge of the process	Watson, 2002
Attack "low hanging" fruit	Smith, 2003; Bicheno, 2001
Involve first-hand observation of target area	Smith, 2003; Vasilash, 1993
Kaizen events used in non-manufacturing areas (e.g., office Kaizen events)	Womack & Jones, 1996a; Sheridan, 1997b; Bradley & Willett, 2004; Melnyk et al., 1998

12-14 members on Kaizen event team	Cuscela, 1998
6 – 10 members on Kaizen event team	McNichols et al., 1999; Martin, 2004; Vasilash, 1993
Training can be provided before the formal start of the event (e.g., offline)	McNichols et al., 1999; Bicheno, 2001
Including process documentation (VSM, process flowcharts, videotapes of the process, current state data, etc.) as input to Kaizen event	Minton, 1998; McNichols et al., 1999; Martin, 2004; Bradley & Willett, 2004; Bicheno, 2001
Notifying employees in adjoining work areas before the start of the Kaizen event	McNichols et al., 1999
Having support personnel (maintenance, engineering, etc.) "on call" during the event, to provide support as needed (e.g., moving equipment overnight)	McNichols et al., 1999; Martin, 2004; Sheridan, 1997b; Bradley & Willett, 2004; Bicheno, 2001; Adams et al., 1997
Training work area in employees in the new process is part of the Kaizen event	Martin, 2004
Concrete, measurable goals	Martin, 2004; Bradley & Willett, 2004; Vasilash, 1993; Melnyk et al., 1998
Black Belts assigned to Kaizen event teams (for Lean-Six Sigma programs)	Sheridan, 2000b
Including benchmarking partners or other external non-supply chain parties on the Kaizen event team	McNichols et al., 1999; Sheridan, 1997b; Vasilash, 1993
Team celebration at the end of the event	Martin, 2004
Organization-wide communication of Kaizen event results	Martin, 2004
Team controls starting and stopping times of Kaizen event activities (often long days 12-14 hrs)	Sheridan, 1997b; Vasilash, 1993
Well-defined and thorough event planning activities (adequate preparation)	Sheridan, 1997b; Bradley & Willett, 2004
Importance of buy-in from employees in work area	Sheridan, 1997b
Keep line running during Kaizen event (important for team to observe a running line)	Sheridan, 1997b
Including target area supervisor on Kaizen event team	Patton, 1997
Use of a "Kaizen office," including full-time coordinators/facilitators	"Keys to Success," 1997; Bicheno, 2001
Total alignment of organizational procedures and policies with Kaizen event program	"Keys to Success," 1997
Stopping production in target area during the Kaizen event	Bradley & Willett, 2004
Including people from all functions required to implement/sustain results on the Kaizen event team	Bradley & Willett, 2004; Vasilash, 1993; Adams et al., 1997
At least one member of Kaizen event team experienced enough in tool(s) to teach others	Bradley & Willett, 2004
At least one member of Kaizen event team keeps the team "on track" (focused)	Bradley & Willett, 2004; Vasilash, 1993
Team should not be too rigid about sticking to formal methodology	Bradley & Willett, 2004
Avoid including people from competing plants or functions on the Kaizen event team	Bradley & Willett, 2004

Preference given to Kaizen events that require simple, well-known tools versus more complex tools	Bradley & Willett, 2004
Cycles of solution refinement during Kaizen event	Bradley & Willett, 2004; Bicheno, 2001; Melnyk et al., 1998
Including people from all production shifts in Kaizen event team	Vasilash, 1993
Teams not punished for failing to meet improvement goals (just asked to understand why)	Vasilash, 1993; Adams et al., 1997; Melnyk et al., 1998
Involving everyone on the Kaizen event team in the solution process	Vasilash, 1993
Including ½ day of training at the start of the event (training in tools, kaizen philosophy, etc.)	Vasilash, 1993; Melnyk et al., 1998
Including ergonomics training as part of Kaizen event training	Wilson, 2005
Combining Kaizen events with other improvement approaches	Bicheno, 2001
Including "team-building" exercises as part of Kaizen event training	Bicheno, 2001
Making sure that each participant has thorough knowledge of the "seven wastes" prior to team activities	Bicheno, 2001
Making each team member responsible for implementing at least one improvement idea	Bicheno, 2001
Kaizen event team members from work area encouraged to discuss event activities and changes with others in the work area during the event (to create buy-in)	Bicheno, 2001
Using a sequence of Kaizen events (e.g., 5S, SMED, Standard Work) to progressively improvement a given work area	Bicheno, 2001; Melnyk et al., 1998
Informal "floating" team structure	Adams et al., 1997
Rewards and recognition for team after the event (e.g., celebrations)	Adams et al., 1997; Melnyk et al., 1998
Each team member participates in report-out to management	Adams et al., 1997
Output of given Kaizen event is used to determine the next Kaizen event	Adams et al., 1997
Kaizen events are focused on the needs of the external customer (e.g. improving value) versus internal efficiency	Melnyk et al., 1998

APPENDIX B: INITIAL GROUPINGS OF FACTORS FROM KAIZEN EVENT LITERATURE

1. Event Design

a) Duration

- One week or shorter (*LeBlanc, 1999; Oakeson, 1997; Vasilash, 1997; Drickhamer, 2004b; Watson, 2002; Smith, 2003; Cuscela, 1998; McNichols et al., 1999; Martin, 2004; Sheridan, 1997b; Patton, 1997; Bradley & Willett, 2004; Vasilash, 1993; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998*)
- Two weeks or shorter (*Minton, 1998; Demers, 2002*)

b) Team Composition

- Team size
    - 3 – 5 people (*Rusiniak, 1996*)
    - 6 – 10 people (*McNichols et al., 1999; Martin, 2004; Vasilash, 1993*)
    - 10 – 12 people (*LeBlanc, 1999; Demers, 2002; Watson, 2002*)
    - 12 – 13 people (*Cuscela, 1998*)
- Use cross-functional teams (*LeBlanc, 1999; Drickhamer, 2004b; Rusiniak, 1996; Demers, 2002; Smith, 2003; Cuscela, 1998; McNichols et al., 1999; Martin, 2004; Sheridan, 1997b; Vasilash, 1993; Vasilash, 1993; Adams et al., 1997; Melnyk et al., 1998*)
    - Informal "floating" team structure (Adams et al., 1997)
    - Team members volunteer to participate (Watson, 2002; Adams et al., 1997)
    - Including "fresh eyes" (people with no prior knowledge of the target area) on the team (LeBlanc, 1999; Vasilash, 1997; Kleinsasser, 2003; Minton, 1998; Cuscela, 1998; McNichols et al., 1999; Martin, 2004; Bradley & Willett, 2004; Melnyk et al., 1998)
    - Including people from the work area on the Kaizen event team (Redding, 1996; Minton, 1998; Womack & Jones, 1996a; Martin, 2004; Sheridan, 1997b; Bradley & Willett, 2004; Vasilash, 1993; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998)
    - Including outside consultants on the Kaizen event team (Oakeson, 1997; Bicheno, 2001)
    - Including managers and supervisors on the Kaizen event team (Oakeson, 1997; "Keys to Success," 1997; Vasilash, 1993; Bicheno, 2001)
    - Including customers on the Kaizen event team (Hasek, 2000; Vasilash, 1997; McNichols et al., 1999; Vasilash, 1993; Adams et al., 1997; Melnyk et al., 1998)
    - Including suppliers on the Kaizen event team (Vasilash, 1997; McNichols et al., 1999; Vasilash, 1993; Adams et al., 1997; Melnyk et al., 1998)
    - Including only one employee per department on the Kaizen event team (except for the department being blitzed), to avoid over-burdening any department (Minton, 1998)
    - Each team member has specific knowledge of the process (Watson, 2002)
    - Black Belts assigned to Kaizen event teams (for Lean-Six Sigma programs) (Sheridan, 2000b)
    - Including benchmarking partners or other external non-supply chain parties on the Kaizen event team (McNichols et al., 1999; Sheridan, 1997b; Vasilash, 1993)
    - Including target area supervisor on Kaizen event team (Patton, 1997)
    - Including people from all functions required to implement/sustain results on the Kaizen event team (Bradley & Willett, 2004; Vasilash, 1993; Adams et al., 1997)
    - At least one member of Kaizen event team experienced enough in tool(s) to teach others (Bradley & Willett, 2004)
    - Avoid including people from competing plants or functions on the Kaizen event team (Bradley & Willett, 2004)
    - Including people from all production shifts in Kaizen event team (Vasilash, 1993)

c) Team Authority

- Teams have implementation authority (*LeBlanc, 1999; Oakeson, 1997; Minton, 1998; Martin, 2004; Sheridan, 1997b; Bradley & Willett, 2004; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998*)
- Team controls starting and stopping times of Kaizen event activities (often long days 12-14 hrs) (*Sheridan, 1997b; Vasilash, 1993*)

d) Problem Scope • Require a standard, reliable target process/work area as input (*LeBlanc, 1999; Bradley & Willett, 2004*) • Requires a well-defined problem statement as input (*Rusiniak, 1996; Adams et al., 1997*) • Avoid problems that are too big and/or emotionally involved (*Rusiniak, 1996;* Sheridan, 1997b) • Preference given to Kaizen events that require simple, well-known tools versus more complex tools (*Bradley & Willett, 2004*) e) Event Goals • Linked to organizational strategy (*LeBlanc, 1999; "Keys to Success," 1997; Melnyk et al., 1998*) • Challenging (stretch) goals (*LeBlanc, 1999; Minton, 1998; Rusiniak, 1996; Cuscela, 1998; Bradley & Willett, 2004; Bicheno, 2001*) • Focused – on a specific process, product, or problem (*Minton, 1998; Drickhamer, 2004b; Martin, 2004; Sheridan, 1997b; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998*) • Used to implement lean manufacturing (*Vasilash, 1997*) • Concrete, measurable goals (*Martin, 2004; Bradley & Willett, 2004; Vasilash, 1993; Melnyk et al., 1998*) • Kaizen events are focused on the needs of the external customer (e.g. improving value) versus internal efficiency (*Melnyk et al., 1998*)
2. Event Planning • Well-defined and thorough event planning activities (adequate preparation) (*Sheridan, 1997b; Bradley & Willett, 2004*) • Including process documentation (VSM, process flowcharts, videotapes of the process, current state data, etc.) as input to Kaizen event (*Minton, 1998; McNichols et al., 1999; Martin, 2004; Bradley & Willett, 2004; Bicheno, 2001*) • Notifying employees in adjoining work areas before the start of the Kaizen event (*McNichols et al., 1999*)
3. Organizational Support a) Management Support/Buy-In (*Bane, 2002; Hasek, 2000; Vasilash, 1997; Rusiniak, 1996; Cuscela, 1998; Martin, 2004; Sheridan, 1997b; "Keys to Success," 1997; Bradley & Willett, 2004; Vasilash, 1993; Bicheno, 2001; Adams et al., 1997*) b) Resource Support • Team members dedicated only to Kaizen event during its duration (*Minton, 1998; McNichols et al., 1999; Martin, 2004; Bradley & Willett, 2004; Bicheno, 2001; Melnyk et al., 1998*) • Having support personnel (maintenance, engineering, etc.) "on call" during the event, to provide support as needed (e.g., moving equipment overnight) (*McNichols et al., 1999; Martin, 2004; Sheridan, 1997b; Bradley & Willett, 2004; Bicheno, 2001; Adams et al., 1997*) • Low cost solutions (*Purdum, 2004; Cuscela, 1998; Martin, 2004; Sheridan, 1997b; Vasilash, 1993; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998*) • Cost is <u>not</u> a factor (*Minton, 1998*) • Dedicated room for Kaizen event team meetings (*Creswell, 2001*) • Snacks provided to team during Kaizen event (*Creswell, 2001; Adams et al., 1997*) • Use of a "Kaizen office," including full-time coordinators/facilitators *("Keys to Success," 1997; Bicheno, 2001*) • Stopping production in target area during the Kaizen event (*Bradley & Willett, 2004*) c) Rewards/Recognition • Rewards and recognition for team after the event (e.g., celebrations) (*Adams et al., 1997; Melnyk et al., 1998*) • Team celebration at the end of the event (*Martin, 2004)* d) Communication • Importance of buy-in from employees in work area (*Sheridan, 1997b*) • Kaizen event team members from work area encouraged to discuss event activities and changes with others in the work area during the event (to create buy-in) (*Bicheno, 2001*)

e) Organizational Policies/Procedures
- "No layoffs" policy (*Redding, 1996; Vasilash, 1997; Creswell, 2001; "Winning with Kaizen," 2002; Womack & Jones, 1996a; "Keys to Success," 1997; Bradley & Willett, 2004; Melnyk et al., 1998*)
- Kaizen event team members from work area encouraged to discuss event activities and changes with others in the work area during the event (to create buy-in) (*Bicheno, 2001*)
- Organization-wide commitment to change (*Redding, 1996*)
- Total alignment of organizational procedures and policies with Kaizen event program ("Keys to Success," 1997)

4. Training
- Less than two hours of formal training provided to team (*Minton, 1998; McNichols et al., 1999)*
- Including ½ day of training at the start of the event (training in tools, kaizen philosophy, etc.) (*Vasilash, 1993; Melnyk et al., 1998*)
- Facilitators provide "short courses" on topics "on the spot" if a team gets stuck (*Minton, 1998*)
- Team members who aren't from the process get training in the process and may even work in the production line for a few days before the Kaizen event (*Minton, 1998*)
- Including ergonomics training as part of Kaizen event training (*Wilson, 2005*)
- Including "team-building" exercises as part of Kaizen event training (*Bicheno, 2001*)
- Making sure that each participant has thorough knowledge of the "seven wastes" prior to team activities (*Bicheno, 2001*)
- Training can be provided before the formal start of the event (e.g., offline) (*McNichols et al., 1999; Bicheno, 2001*)

5. Systematic Use of Kaizen Events
- Spacing out events (e.g., only 1 event per quarter) (*Taninecz, 1997)*
- Concurrent Kaizen events (*Vasilash, 1997; Watson, 2002; Cuscela, 1998; Bradley & Willett, 2004; Adams et al., 1997*)
- Targeted at areas that can provide a "big win" (big impact on organization) (*Minton, 1998; Cuscela, 1998; Martin, 2004; Sheridan, 1997b; "Keys to Success," 1997; Bradley & Willett, 2004; Melnyk et al., 1998*)
- Repeat Kaizen events in a given work area (*"Winning with Kaizen," 2002; Purdum, 2004; Womack & Jones, 1996a; McNichols et al., 1999; Sheridan, 1997b; Bradley & Willett, 2004; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998*)
- Can be held based on employee suggestions for improvement (*Jusko, 2004; Watson, 2002*)
- Kaizen events used in non-manufacturing areas (e.g., office Kaizen events) (*Womack & Jones, 1996a; Sheridan, 1997b; Bradley & Willett, 2004; Melnyk et al., 1998*)
- Combining Kaizen events with other improvement approaches (*Bicheno, 2001*)
- Using a sequence of Kaizen events (e.g., 5S, SMED, Standard Work) to progressively improvement a given work area (*Bicheno, 2001; Melnyk et al., 1998*)
- Attack "low hanging fruit" (*Smith, 2003; Bicheno, 2001*)
- Output of given Kaizen event is used to determine the next Kaizen event (*Adams et al., 1997*)

6. Event Process

a) Action Orientation (*LeBlanc, 1999; Redding, 1996; Smith, 2003; Martin, 2004; Sheridan, 1997b; Patton, 1997; Vasilash, 1993; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998*)
- Involve first-hand observation of target area (*Smith, 2003; Vasilash, 1993*)
- Keep line running during Kaizen event (important for team to observe a running line) (*Sheridan, 1997b*)
- Cycles of solution refinement during Kaizen event (*Bradley & Willett, 2004; Bicheno, 2001; Melnyk et al., 1998*)
- Training work area in employees in the new process is part of the Kaizen event (*Martin, 2004*)

b) Problem Solving Tools/Techniques
- Videotapes of setups (*Minton, 1998; Bradley & Willett, 2004*)
- Brainstorming (*Minton, 1998; Watson, 2002; Martin, 2004; Bradley & Willett, 2004; Vasilash, 1993)*

- Avoid preconceived solutions (*Rusiniak, 1996; Bradley & Willett, 2004*)
- Seek improvement, not optimization (*Rusiniak, 1996; Vasilash, 1993*)
- Question the current process – ask why things are done the way they are (*Watson, 2002; Minton, 1998*)
- Team should not be too rigid about sticking to formal methodology (*Bradley & Willett, 2004*)

c) Team Coordination
- At least one member of Kaizen event team keeps the team "on track" (focused) (*Bradley & Willett, 2004; Vasilash, 1993*)
- Use of subteams (*Minton, 1998; McNichols et al., 1999; Sheridan, 1997b; Bicheno, 2001*)
- Use of a Kaizen newspaper (*"Winning with Kaizen," 2002; McNichols et al., 1999; Martin, 2004; Bradley & Willett, 2004; Melnyk et al., 1998*)

d) Participation
- Involving everyone on the Kaizen event team in the solution process (*Vasilash, 1993*)
- Making each team member responsible for implementing at least one improvement idea (*Bicheno, 2001*)
- Each team member participates in report-out to management (*Adams et al., 1997*)

APPENDIX C: CATEGORIES OF FACTORS FROM KAIZEN EVENT LITERATURE

(based on initial 33 sources reviewed)

1. **Task** -- Task Design Factors (Cohen & Bailey, 1997); Project Level Antecedents (Nicolini, 2002); Factors Related to Project (Belassi & Tukel, 1996)

    a) Event Duration
        - One week or shorter (*LeBlanc, 1999; Oakeson, 1997; Vasilash, 1997; Drickhamer, 2004b; Watson, 2002; Smith, 2003; Cuscela, 1998; McNichols et al., 1999; Martin, 2004; Sheridan, 1997b; Patton, 1997; Bradley & Willett, 2004; Vasilash, 1993; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998*)
        - Two weeks or shorter (*Minton, 1998; Demers, 2002*)

    b) Team Authority
        - Teams have implementation authority (*LeBlanc, 1999; Oakeson, 1997; Minton, 1998; Martin, 2004; Sheridan, 1997b; Bradley & Willett, 2004; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998*)
        - Team controls starting and stopping times of Kaizen event activities (often long days 12-14 hrs) (*Sheridan, 1997b; Vasilash, 1993*)

    c) Problem Scope
        - Require a standard, reliable target process/work area as input (*LeBlanc, 1999; Bradley & Willett, 2004*)
        - Requires a well-defined problem statement as input (*Rusiniak, 1996; Adams et al., 1997*)
        - Avoid problems that are too big and/or emotionally involved (*Rusiniak, 1996;* Sheridan, 1997b)
        - Preference given to Kaizen events that require simple, well-known tools versus more complex tools (*Bradley & Willett, 2004*)

    d) Event Goals
        - Linked to organizational strategy (*LeBlanc, 1999; "Keys to Success," 1997; Melnyk et al., 1998*)
        - Challenging (stretch) goals (*LeBlanc, 1999; Minton, 1998; Rusiniak, 1996; Cuscela, 1998; Bradley & Willett, 2004; Bicheno, 2001*)
        - Focused – on a specific process, product, or problem (*Minton, 1998; Drickhamer, 2004b; Martin, 2004; Sheridan, 1997b; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998*)
        - Used to implement lean manufacturing (*Vasilash, 1997*)
        - Concrete, measurable goals (*Martin, 2004; Bradley & Willett, 2004; Vasilash, 1993; Melnyk et al., 1998*)
        - Kaizen events are focused on the needs of the external customer (e.g. improving value) versus internal efficiency (*Melnyk et al., 1998*)

2. **Team** -- Group Composition Factors (Cohen & Bailey, 1997); Project Level Antecedents (Nicolini, 2002); Factors Related to the Project Team (Belassi & Tukel, 1996)

    a) Team size
        - 3 – 5 people (Rusiniak, 1996)
        - 6 – 10 people (McNichols et al., 1999; Martin, 2004; Vasilash, 1993)
        - 10 – 12 people (LeBlanc, 1999; Demers, 2002; Watson, 2002)
        - 12 – 13 people (Cuscela, 1998)

    b) Use of Cross-Functional Teams (*LeBlanc, 1999; Drickhamer, 2004b; Rusiniak, 1996; Demers, 2002; Smith, 2003; Cuscela, 1998; McNichols et al., 1999; Martin, 2004; Sheridan, 1997b; Vasilash, 1993; Vasilash, 1993; Adams et al., 1997; Melnyk et al., 1998*)
        - Informal "floating" team structure (*Adams et al., 1997*)
        - Team members volunteer to participate (*Watson, 2002; Adams et al., 1997*)
        - Including "fresh eyes" (people with no prior knowledge of the target area) on the team (*LeBlanc, 1999; Vasilash, 1997; Kleinsasser, 2003; Minton, 1998; Cuscela, 1998; McNichols et al., 1999; Martin, 2004; Bradley & Willett, 2004; Melnyk et al., 1998*)
        - Including people from the work area on the Kaizen event team (*Redding, 1996; Minton, 1998; Womack & Jones, 1996a; Martin, 2004; Sheridan, 1997b; Bradley & Willett, 2004; Vasilash, 1993; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998*)

- Including outside consultants on the Kaizen event team (*Oakeson, 1997; Bicheno, 2001*)
- Including managers and supervisors on the Kaizen event team (*Oakeson, 1997; "Keys to Success," 1997; Vasilash, 1993; Bicheno, 2001*)
- Including customers on the Kaizen event team (*Hasek, 2000; Vasilash, 1997; McNichols et al., 1999; Vasilash, 1993; Adams et al., 1997; Melnyk et al., 1998*)
- Including suppliers on the Kaizen event team (*Vasilash, 1997; McNichols et al., 1999; Vasilash, 1993; Adams et al., 1997; Melnyk et al., 1998*)
- Including only one employee per department on the Kaizen event team (except for the department being blitzed), to avoid over-burdening any department (*Minton, 1998*)
- Each team member has specific knowledge of the process (*Watson, 2002*)
- Black Belts assigned to Kaizen event teams (for Lean-Six Sigma programs) (*Sheridan, 2000b*)
- Including benchmarking partners or other external non-supply chain parties on the Kaizen event team (*McNichols et al., 1999; Sheridan, 1997b; Vasilash, 1993*)
- Including target area supervisor on Kaizen event team (*Patton, 1997*)
- Including people from all functions required to implement/sustain results on the Kaizen event team (*Bradley & Willett, 2004; Vasilash, 1993; Adams et al., 1997*)
- At least one member of Kaizen event team experienced enough in tool(s) to teach others (*Bradley & Willett, 2004*)
- Avoid including people from competing plants or functions on the Kaizen event team (*Bradley & Willett, 2004*)
- Including people from all production shifts in Kaizen event team (*Vasilash, 1993*)

3. **Organization** -- Organizational Context Factors (Cohen & Bailey, 1997); Project Level Antecedents (Nicolini, 2002); Factors Related to the Organization (Belassi & Tukel, 1996)

   a) Management Support/Buy-In (*Bane, 2002; Hasek, 2000; Vasilash, 1997; Rusiniak, 1996; Cuscela, 1998; Martin, 2004; Sheridan, 1997b; "Keys to Success," 1997; Bradley & Willett, 2004; Vasilash, 1993; Bicheno, 2001; Adams et al., 1997*)
   b) Resource Support
      - Team members dedicated only to Kaizen event during its duration (*Minton, 1998; McNichols et al., 1999; Martin, 2004; Bradley & Willett, 2004; Bicheno, 2001; Melnyk et al., 1998*)
      - Having support personnel (maintenance, engineering, etc.) "on call" during the event, to provide support as needed (e.g., moving equipment overnight) (*McNichols et al., 1999; Martin, 2004; Sheridan, 1997b; Bradley & Willett, 2004; Bicheno, 2001; Adams et al., 1997*)
      - Low cost solutions (*Purdum, 2004; Cuscela, 1998; Martin, 2004; Sheridan, 1997b; Vasilash, 1993; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998*)
      - Cost is <u>not</u> a factor (*Minton, 1998*)
      - Dedicated room for Kaizen event team meetings (*Creswell, 2001*)
      - Snacks provided to team during Kaizen event (*Creswell, 2001; Adams et al., 1997*)
      - Use of a "Kaizen office," including full-time coordinators/facilitators *("Keys to Success," 1997; Bicheno, 2001*)
      - Stopping production in target area during the Kaizen event (*Bradley & Willett, 2004*)
   c) Rewards/Recognition
      - Rewards and recognition for team after the event (e.g., celebrations) (*Adams et al., 1997; Melnyk et al., 1998*)
      - Team celebration at the end of the event (*Martin, 2004)*
   d) Communication
      - Importance of buy-in from employees in work area (*Sheridan, 1997b*)
      - Kaizen event team members from work area encouraged to discuss event activities and changes with others in the work area during the event (to create buy-in) (*Bicheno, 2001*)
   e) Event Planning Process
      - Well-defined and thorough event planning activities (adequate preparation) (*Sheridan, 1997b; Bradley & Willett, 2004*)
      - Including process documentation (VSM, process flowcharts, videotapes of the process, current state

data, etc.) as input to Kaizen event (*Minton, 1998; McNichols et al., 1999; Martin, 2004; Bradley & Willett, 2004; Bicheno, 2001*)
- Notifying employees in adjoining work areas before the start of the Kaizen event (*McNichols et al., 1999*)

f) Training
- Less than two hours of formal training provided to team (*Minton, 1998; McNichols et al., 1999)*
- Including ½ day of training at the start of the event (training in tools, kaizen philosophy, etc.) (*Vasilash, 1993; Melnyk et al., 1998*)
- Facilitators provide "short courses" on topics "on the spot" if a team gets stuck (*Minton, 1998*)
- Team members who aren't from the process get training in the process and may even work in the production line for a few days before the Kaizen event (*Minton, 1998*)
- Including ergonomics training as part of Kaizen event training (*Wilson, 2005*)
- Including "team-building" exercises as part of Kaizen event training (*Bicheno, 2001*)
- Making sure that each participant has thorough knowledge of the "seven wastes" prior to team activities (*Bicheno, 2001*)
- Training can be provided before the formal start of the event (e.g., offline) (*McNichols et al., 1999; Bicheno, 2001*)

4. **Event Process** -- Internal Process Factors (Cohen & Bailey, 1997); Processes (Nicolini, 2002); Project Manager's Performance on the Job (Belassi & Tukel, 1996)

a) Action Orientation (*LeBlanc, 1999; Redding, 1996; Smith, 2003; Martin, 2004; Sheridan, 1997b; Patton, 1997; Vasilash, 1993; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998*)
- Involve first-hand observation of target area (*Smith, 2003; Vasilash, 1993*)
- Keep line running during Kaizen event (important for team to observe a running line) (*Sheridan, 1997b*)
- Cycles of solution refinement during Kaizen event (*Bradley & Willett, 2004; Bicheno, 2001; Melnyk et al., 1998*)
- Training work area in employees in the new process is part of the Kaizen event (*Martin, 2004*)

b) Problem Solving Tools/Techniques
- Videotapes of setups (*Minton, 1998; Bradley & Willett, 2004*)
- Brainstorming (*Minton, 1998; Watson, 2002; Martin, 2004; Bradley & Willett, 2004; Vasilash, 1993)*
- Avoid preconceived solutions (*Rusiniak, 1996; Bradley & Willett, 2004*)
- Seek improvement, not optimization (*Rusiniak, 1996; Vasilash, 1993*)
- Question the current process – ask why things are done the way they are (*Watson, 2002; Minton, 1998*)
- Team should not be too rigid about sticking to formal methodology (*Bradley & Willett, 2004*)

c) Team Coordination
- At least one member of Kaizen event team keeps the team "on track" (focused) (*Bradley & Willett, 2004; Vasilash, 1993*)
- Use of subteams (*Minton, 1998; McNichols et al., 1999; Sheridan, 1997b; Bicheno, 2001*)
- Use of a Kaizen newspaper (*"Winning with Kaizen," 2002; McNichols et al., 1999; Martin, 2004; Bradley & Willett, 2004; Melnyk et al., 1998*)

d) Participation
- Involving everyone on the Kaizen event team in the solution process (*Vasilash, 1993*)
- Making each team member responsible for implementing at least one improvement idea (*Bicheno, 2001*)
- Each team member participates in report-out to management (*Adams et al., 1997*)

5. **Broader Context** (Kaizen Event Program Characteristics)

a) Kaizen Event Deployment
- Spacing out events (e.g., only 1 event per quarter) (*Taninecz, 1997)*
- Concurrent Kaizen events (*Vasilash, 1997; Watson, 2002; Cuscela, 1998; Bradley & Willett, 2004; Adams et al., 1997*)

- Targeted at areas that can provide a "big win" (big impact on organization) (*Minton, 1998; Cuscela, 1998; Martin, 2004; Sheridan, 1997b; "Keys to Success," 1997; Bradley & Willett, 2004; Melnyk et al., 1998*)
- Repeat Kaizen events in a given work area (*"Winning with Kaizen," 2002; Purdum, 2004; Womack & Jones, 1996a; McNichols et al., 1999; Sheridan, 1997b; Bradley & Willett, 2004; Bicheno, 2001; Adams et al., 1997; Melnyk et al., 1998*)
- Can be held based on employee suggestions for improvement (*Jusko, 2004; Watson, 2002*)
- Kaizen events used in non-manufacturing areas (e.g., office Kaizen events) (*Womack & Jones, 1996a; Sheridan, 1997b; Bradley & Willett, 2004; Melnyk et al., 1998*)
- Combining Kaizen events with other improvement approaches (*Bicheno, 2001*)
- Using a sequence of Kaizen events (e.g., 5S, SMED, Standard Work) to progressively improvement a given work area (*Bicheno, 2001; Melnyk et al., 1998*)
- Attack "low hanging fruit" (*Smith, 2003; Bicheno, 2001*)
- Output of given Kaizen event is used to determine the next Kaizen event (*Adams et al., 1997*)

b) Organizational Policies/Procedures
- "No layoffs" policy (*Redding, 1996; Vasilash, 1997; Creswell, 2001; "Winning with Kaizen," 2002; Womack & Jones, 1996a; "Keys to Success," 1997; Bradley & Willett, 2004; Melnyk et al., 1998*)
- Kaizen event team members from work area encouraged to discuss event activities and changes with others in the work area during the event (to create buy-in) (*Bicheno, 2001*)
- Organization-wide commitment to change (*Redding, 1996*)
- Total alignment of organizational procedures and policies with Kaizen event program ("Keys to Success," 1997)

APPENDIX D: EXAMPLE KAIZEN EVENT ANNOUNCEMENT

Kaizen Event Announcement
Standard Work: Cell A
October 4th - 8th

TO: Team Member 1, Team Member 2, Team Member 3, Team Member 4, Team Member 5, Team Member 6

FROM: Plant Manager, Kaizen Event Facilitator, Kaizen Event Team Leader

SUBJECT: Standard Work: Cell A

CC: Company A Ops Staff, Company A Kaizen Event Coordinator, Cell A Supervisor, Cell A

DATE: September 1st

Congratulations! You have been selected to participate in an upcoming Kaizen event for Cell A. This is a great opportunity to make a major improvement in quality and delivery performance. We will focus on combining the sub-process 1 and sub-process 2 cells to create a complete process cell.

MEETING TIME/PLACE:
Training for this event will start on Friday, October 1, from 8:00 a.m. to 4:00 p.m. Training will be held in the Training room. Kick-off for this event will start Monday, October 4, at 8:00 a.m. in the Training room. This event will continue at Facility A through Friday, October 8. On Friday, October 8 at 11:00 a.m., we will report out to the management team. For those on CC, please plan to attend the report out. The wrap-up is very important.

BUSINESS ISSUE:
This cell has consistently > 93% OTD and very low external warranty returns. However, the most significant barrier to further growth is the relatively long lead times (4-6 weeks) for most items. Implementing one-piece flow will improve the flexibility of the cell, enabling us to increase capacity while significantly reducing lead times and reducing the need for cell associates to work overtime. Additionally, this will reduce WIP and inventory.

PROPOSAL:
Create a high-powered, cross-functional team for a 4 1/2 - day kaizen to implement one-piece flow in Cell A.

TEAM GOALS & OBJECTIVES:
- Implement one piece flow
- Increase cell capacity
- Improve cell flexibility and responsiveness.
- Create Standard Work documents and daily management process (including metrics)
- Achieve a 4S rating and develop a plan for achieving and maintaining a 5S rating.

HOW WILL SUCCESS BE MEASURED:
- One-piece flow established
- Increase throughput by 35%.
- Reduce lead-time by 75%.
- Documents created and playbooks in place. Standard WIP calculated & in place. At 90 days, 15% reduction in inventory dollars.
- 4S rating achieved after audit.

OTHER NOTES:

Day	Meeting Times	Meeting Purpose	Attendees	Location
Friday October 1st	8:00am - 12:00pm	Training	Kaizen Team	Training Room
Monday - Thursday October 4 – 7	8:00am -???	Working Session	Kaizen Team	Training Room
Friday October 8th	8:00am -11:00am	Working Session & Report Out	Kaizen Team; Mgmt. & CC	Training Room

APPENDIX E: PILOT VERSION OF KICKOFF SURVEY

Kickoff Survey

Hello Kaizen Team Member,

Your help is needed on this important research. This survey is part of a research project sponsored by the National Science Foundation. The research studies the effects of Kaizen events and what makes them successful. Your company is one of the few companies chosen for the research and will get first access to the results. Your company will use the results to design better Kaizen events and better support its Kaizen event teams.

This survey asks for your opinions on the goals of your Kaizen event team. The survey is short and should only take about 10 minutes to complete.

Participation in this survey is voluntary. You are not asked to provide your name, just the name of your Kaizen event team. The survey confidential and the privacy of your answers will be protected. No one at your company will see your individual answers.

If you wish to participate, please answer the questions on the next page (page 2). The questions ask you to describe how much you agree or disagree with statements about the goals of your team. If a question does not seem to be applicable, please answer “disagree” or “strongly disagree.”

In the survey, “Kaizen event” refers to your Kaizen event team and not to any other teams working at the same time. “Work area” refers to the manufacturing area that is the focus of your Kaizen event.

Thank you for your help in this important research! If you have any questions or comments, please contact [insert name of University contact, e-mail, and phone number].

Kaizen Team Name ________________________________	Strongly Disagree 1	Disagree 2	Tend to Disagree 3	Tend to Agree 4	Agree 5	Strongly Agree 6
Circle the response that BEST describes your opinion.						
1. In general, members of our [insert name of Kaizen event] team believe that this Kaizen event is needed.	1	2	3	4	5	6
2. It will take a lot of skill to achieve our [insert name of Kaizen event] team's improvement goals.	1	2	3	4	5	6
3. Our entire [insert name of Kaizen event] team understands our goals	1	2	3	4	5	6
4. In general, members of our [insert name of Kaizen event] team believe in the value of this Kaizen event.	1	2	3	4	5	6
5. Our goals clearly define what is expected of our [insert name of Kaizen event] team	1	2	3	4	5	6
6. Meeting our [insert name of Kaizen event] team's improvement goals will be tough	1	2	3	4	5	6
7. Most of our [insert name of Kaizen event] team members think that this Kaizen event is a good strategy for this work area.	1	2	3	4	5	6
8. Most of our [insert name of Kaizen event] team members think that things would be better without this Kaizen event.	1	2	3	4	5	6
9. Our [insert name of Kaizen event] team's improvement goals are difficult	1	2	3	4	5	6
10. Most of our [insert name of Kaizen event] team members that this Kaizen event will serve an important purpose.	1	2	3	4	5	6
11. It will be hard to improve this work area enough to achieve our [insert name of Kaizen event] team's goals	1	2	3	4	5	6
12. Our [insert name of Kaizen event] team has clearly defined goals	1	2	3	4	5	6
13. In general, members of our [insert name of Kaizen event] team think that it is a mistake to hold this Kaizen event	1	2	3	4	5	6
14. The performance targets our [insert name of Kaizen event] team must achieve to fulfill our goals are clear	1	2	3	4	5	6

15. How many Kaizen events total have you participated in? ***(Fill in the blank)*** ________________________

16. Which functional area most closely describes your current job? ***(Circle one number)***
    1. OPERATIONS (ex. Production Associate, Maintenance Technician, etc.)
    2. NON-OPERATIONS (ex. Engineer, Marketing Associate, Manager, etc.)

Thank you for your participation! If you have any other comments about your experience with Kaizen events, please include them on the back of this page.

APPENDIX F: FINAL VERSION OF KICKOFF SURVEY

Kickoff Survey

Hello [insert name of Kaizen event] Team Member,

Your help is needed on this important research. This survey is part of a research project sponsored by the National Science Foundation. The research studies the effects of Kaizen events and what makes them successful. Your company is one of the few companies chosen for the research and will get first access to the results. Your company will use the results to design better Kaizen events and better support its Kaizen event teams.

This survey asks for your opinions on the goals of your Kaizen event team (the [insert name of Kaizen event] team). The survey is short and should only take about 10 minutes to complete.

The survey asks you to choose a unique survey code that will allow the researchers to perform analysis on the survey results. This code is very important. It will not be used to identify you, since only you will know what your code is. If you have already used a survey code on previous surveys for this research, please use the same code. If you have not completed any previous surveys for this research, please use one of the next three options to choose your survey code:

Option 1: The first four letters of your mother's maiden name [ex. Brown = "brow"]
Option 2: The month and day of your birthday [ex. January 1 = "0101"]
Option 3: The first four letters of your pet or child's name [ex. Dolly = "doll"]

Write the survey code you select on the top of the next page (page 2), after the words "Survey Code." Please remember the code you select, since it will be used for future surveys. **Do not include your name.** The survey is confidential and the privacy of your answers will be protected. No one at your company will see your individual answers.

Participation in this survey is voluntary. If you wish to participate, please answer the questions on the next page (page 2) and return the survey to the envelope provided. The questions ask you to describe how much you agree or disagree with statements about the goals of your team. You may decline to answer any question(s) you choose. If a question does not seem to apply to your team, it may be that you "disagree" or "strongly disagree" with the statement.

In the survey, "Kaizen event" refers to your Kaizen event team and not to any other teams working at the same time. "Work area" refers to the manufacturing area that is the focus of your Kaizen event.

Thank you for your help in this important research! If you have any questions or comments, please contact [insert name of University contact, e-mail, and phone number]. If you have questions about your rights as a research participant, please contact the Virginia Tech Institutional Review Board (IRB) Human Protections Administrator, [insert name, e-mail and phone number of university Institutional Review Board (IRB) Human Protections Administrator].

Kaizen Team Name: [insert name of Kaizen event] Survey Code:	Strongly Disagree 1	Disagree 2	Tend to Disagree 3	Tend to Agree 4	Agree 5	Strongly Agree 6
Circle the response that BEST describes your opinion.						
1. Most of our [insert name of Kaizen event] team members think that this Kaizen event is a good strategy for this work area.	1	2	3	4	5	6
2. It will take a lot of effort to achieve our [insert name of Kaizen event] team's goals.	1	2	3	4	5	6
3. The performance targets our [insert name of Kaizen event] team must achieve to fulfill our goals are clear.	1	2	3	4	5	6
4. In general, members of our [insert name of Kaizen event] team believe in the value of this Kaizen event.	1	2	3	4	5	6
5. Our [insert name of Kaizen event] team has enough time to achieve our goals.	1	2	3	4	5	6
6. Most of our [insert name of Kaizen event] team members think that things will be better with this Kaizen event.	1	2	3	4	5	6
7. It will be hard to improve this work area enough to achieve our [insert name of Kaizen event] team's goals.	1	2	3	4	5	6
8. It will take a lot of thought to achieve our [insert name of Kaizen event] team's goals.	1	2	3	4	5	6
9. In general, members of our [insert name of Kaizen event] team believe that this Kaizen event is needed.	1	2	3	4	5	6
10. Our [insert name of Kaizen event] team has clearly defined goals.	1	2	3	4	5	6
11. Our [insert name of Kaizen event] team's goals are difficult.	1	2	3	4	5	6
12. Most of our [insert name of Kaizen event] team members believe that this Kaizen event will serve an important purpose.	1	2	3	4	5	6
13. Our entire [insert name of Kaizen event] team understands our goals.	1	2	3	4	5	6
14. Meeting our [insert name of Kaizen event] team's goals will be tough.	1	2	3	4	5	6
15. Our goals clearly define what is expected of our [insert name of Kaizen event] team.	1	2	3	4	5	6
16. In general, members of our [insert name of Kaizen event] team think that it is a mistake to hold this Kaizen event.	1	2	3	4	5	6
17. It will take a lot of skill to achieve our [insert name of Kaizen event] team's goals.	1	2	3	4	5	6

18. Not including this event, how many Kaizen events total have you participated in? ***(Fill in the blank)***

____

19. Which functional area most closely describes your current job? ***(Circle one number)***
    1. OPERATIONS (ex. Production Associate, Maintenance Technician, etc.)
    2. NON-OPERATIONS (ex. Engineer, Marketing Associate, Manager, etc.)

Thank you for your participation! If you have any other comments about your experience with Kaizen events, please include them on the back of this page.

APPENDIX G: PILOT VERSION OF TEAM ACTIVITIES LOG

Team Activities Log

Hello Kaizen Team Member,

Your help is needed on this important research. This form is part of a research project sponsored by the National Science Foundation. The research studies the effects of Kaizen events and what makes them successful. Your company is one of the few companies chosen for the research and will get first access to the results. Your company will use the results to design better Kaizen events and better support its Kaizen event teams.

This form asks you to briefly record the daily activities Kaizen event team. The form is short and should only take about 15 minutes total to complete.

Participation in this research is voluntary and confidential. You are not asked to provide your name, just the name of your Kaizen event team.

If you wish to participate, use the timetable ("Team Activities Log") on page 3 to briefly describe the daily activities of your team. Please write a brief description of each major activity (ex. setup, videotape review, etc) and indicate where the activity took place (meeting room or target work area). In addition, please indicate the tools your team used for each activity (ex. brainstorming, Pareto analysis). Finally, please indicate whether your team was working as a group or in subgroups.

See page 2 for an example of a completed timetable ("Team Activities Log").

At the end of the Kaizen event, please return the completed form to [insert facilitator name] in the envelop provided. Thank you for your help in this important research! If you have any questions or comments, please contact [insert name of University contact, e-mail, and phone number].

.

Example of Completed Team Activities Log

Kaizen Event Team: Alpha Line Standard Work

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6
7:30		Observe current state and collect data using time charts, sphaghetti charts, etc. (work area)				
8:00	Kickoff to introduction goals and objectives of the event. Members of senior management in attendance. (meeting room)		Train operators in new standard work procedures (work area)	Split into subgroups. Group 1 = implementing supermarket quantities (work area). Group 2 = train operators in new standard work procedures and test new procedures using time charts, etc. (work area)		
8:30					Finish report-out presentation and practice presenting (meeting room).	
9:00		Split into subgroups. Group 1 = working on new cell layout ideas (meeting room) Group 2 = working on improving standard work procedures using takt time calculations, seven wastes of lean,etc. (meeting room)	Test new standard work procedures using time charts, etc. (work area)			
9:30	Standard work training (meeting room)					
10:00					Report-out presentation to management (meeting room)	
10:30				Regroup and report progress. Split again into subgroups. Group 1 = apply 5S to work area (work area) Group = work on standard work documentation. (meeting room)		
11:00					Celebration Lunch	
11:30						
12:00	Lunch	Lunch	Lunch	Lunch		
12:30	Standard work training (meeting room)	Regroup and present progress (meeting room)	Split into subgroups. Group 1 = calculating supermarket quantities for cell (meeting room) Group 2 = brainstorming further improvements to standard work using takt time calculations, seven wastes of lean,etc. (meeting room)	Continue working in subgroups. Group 1 = completing 5S and training operators in 5S procedures (work area) Group 2 = complete standard work documentation. (meeting room)		
1:00						
1:30		Team leader in meeting with facilitator (meeting room). Rest of the group moving equipment to implement new cell layout (work area)				
2:00						
2:30				Regroup. Place standard work documentation in work area and train operators in standard work documentation (work area)		
3:00						
3:30		Team leader rejoins group, continuing hooking up equipment (work area)	Regroup and resport progress. Facilitator and member of management responsible for work area in attendance (meeting room)	Meet with facilitator (meeting room)		
4:00				Begin work on report out presentation (meeting room)		
4:30						
5:00						
5:30						
6:00						
6:30						

Team Activities Log – ***Please Use this Page to Briefly Record Your Team's Daily Activities***

Kaizen Event Team: ________________________________

	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6
7:30						
8:00						
8:30						
9:00						
9:30						
10:00						
10:30						
11:00						
11:30						
12:00						
12:30						
1:00						
1:30						
2:00						
2:30						
3:00						
3:30						
4:00						
4:30						
5:00						
5:30						
6:00						
6:30						

APPENDIX H: FINAL VERSION OF TEAM ACTIVITIES LOG

Team Activities Log

Hello [insert name of Kaizen event] Team Member,

Your help is needed on this important research. This form is part of a research project sponsored by the National Science Foundation. The research studies the effects of Kaizen events and what makes them successful. Your company is one of the few companies chosen for the research and will get first access to the results. Your company will use the results to design better Kaizen events and better support its Kaizen event teams.

This form asks you to briefly record the daily activities Kaizen event team (the [insert name of Kaizen event] team). The form is short and should only take about 15 minutes total to complete.

Participation in this research is voluntary and confidential. You are not asked to provide your name, just the name of your Kaizen event team.

If you wish to participate, please use the timetable ("Team Activities Log") on pages 4 – [insert page number based on length of event] to briefly describe the daily activities of the [insert name of Kaizen event] team. The time labels can refer to either am or pm. Please write "am" or "pm" in your descriptions to indicate the timing of your event (see example Logs). Please write a brief description of each major activity (ex. setup, videotape review, etc) and indicate where the activity took place (meeting room or target work area). In addition, please indicate the tools the [insert name of Kaizen event] team used for each activity (ex. brainstorming, Pareto analysis, etc.). Finally, please indicate whether the [insert name of Kaizen event] team was working as a group or in subgroups.

See page 2 for an example of a completed Team Activities Log for a daytime event. See page for an example of a completed Team Activities Log for a nighttime event.

At the end of the [insert name of Kaizen event] event, please return the completed form to [insert facilitator name] in the envelope provided. Thank you for your help in this important research! If you have any questions or comments, please contact [insert name of University contact, e-mail, and phone number]. If you have questions about your rights as a research participant, please contact the Virginia Tech Institutional Review Board (IRB) Human Protections Administrator, [insert name, e-mail and phone number of university Institutional Review Board (IRB) Human Protections Administrator].

Example of Completed Team Activities Log: Daytime Event (Day 1 – Day 3 of a 5 day event)

Kaizen Event Team: Alpha Line Standard Work

	Day 1	Day 2	Day 3
5:00			
5:30			
6:00			
6:30			
7:00			
7:30		AM. Observe current state and collect data using time charts & spaghetti charts (work area)	
8:00	AM. Kickoff to introduce goals and objectives of the event. Members of senior management in attendance. (meeting room)		AM. Train operators in new standard work procedures (work area)
8:30			
9:00		AM. Split into subgroups. Group 1 = working on new cell layout ideas (meeting room) Group 2 = working on improving standard work procedures using takt time calculations and seven wastes of lean (meeting room)	AM. Test new standard work procedures using time charts and observation (work area)
9:30	AM. Standard work training (meeting room)		
10:00			
10:30			
11:00			
11:30			
12:00	PM. Lunch	PM. Lunch	PM. Lunch
12:30	PM. Standard work training (meeting room)	PM. Regroup and present progress (meeting room)	PM. Split into subgroups. Group 1 = calculating supermarket quantities for cell (meeting room) Group 2 = brainstorming further improvements to standard work using takt time calculations, seven wastes of lean, etc. (meeting room)
1:00			
1:30		PM. Team leader in meeting with facilitator (meeting room). Rest of the group moving equipment to implement new cell layout (work area)	
2:00			
2:30			
3:00			
3:30		PM. Team leader rejoins group, continuing hooking up equipment (work area)	PM. Regroup and report progress. Facilitator and member of management responsible for work area in attendance (meeting room)
4:00			
4:30			
5:00			
5:30			
6:00			
6:30			
7:00			

Example of Completed Team Activities Log: Nighttime Event (Day 1 – Day 3 of a 5 day event)

Kaizen Event Team: Machine B TPM

	Day 1	Day 2	Day 3
5:00			
5:30			
6:00			
6:30			
7:00			PM. Team meeting to discuss yesterday's progress and plan for today's work. Developed a new list of improvements to work on today. Assigned items to team members (meeting room)
7:30			
8:00		PM. Team meeting to discuss plans/agenda for the day (meeting room)	
8:30			
9:00		PM. Team members went out to the work area to work on assigned action items (work area)	
9:30			
10:00	PM. Kickoff meeting (introduction to the goals of the event) (meeting room)		
10:30			
11:00			PM. Dinner
11:30	PM. Training (meeting room)		AM. Team members went out to the work area to work on assigned action items (work area)
12:00			
12:30		AM. Dinner	
1:00			
1:30		AM. Worked as a team on more difficult improvements (work area)	
2:00	AM. Lunch/Breakfast		
2:30	AM. Break		
3:00	AM. Training (meeting room)	AM. Painting and labeling (work area)	AM. Painting (work area)
3:30	AM. Pre-TPM Evaluation of Machine B (work area)		AM. Clean-up (work area)
4:00			AM. Team meeting to talk about progress (meeting room)
4:30		AM. Clean-up (work area)	
5:00	AM. Developed a list of items to work on and a action plan for working on them. (meeting room)		
5:30			
6:00			
6:30			
7:00			

Team Activities Log – ***Please Use this Page to Briefly Record the [insert name of Kaizen event] Team's Daily Activities***

Kaizen Event Team: [insert name of Kaizen event]

	Day 1	Day 2	Day 3
5:00			
5:30			
6:00			
6:30			
7:00			
7:30			
8:00			
8:30			
9:00			
9:30			
10:00			
10:30			
11:00			
11:30			
12:00			
12:30			
1:00			
1:30			
2:00			
2:30			
3:00			
3:30			
4:00			
4:30			
5:00			
5:30			
6:00			
6:30			
7:00			

Team Activities Log – ***Please Use this Page to Continue Recording the [insert name of Kaizen event] Team's Daily Activities***

Kaizen Event Team: [insert name of Kaizen event]

	Day 4	Day 5	Day 6
5:00			
5:30			
6:00			
6:30			
7:00			
7:30			
8:00			
8:30			
9:00			
9:30			
10:00			
10:30			
11:00			
11:30			
12:00			
12:30			
1:00			
1:30			
2:00			
2:30			
3:00			
3:30			
4:00			
4:30			
5:00			
5:30			
6:00			
6:30			
7:00			

APPENDIX I: PILOT VERSION OF REPORT OUT SURVEY

Report Out Survey

Hello Kaizen Team Member,

Your help is needed on this important research. This survey is part of a research project sponsored by the National Science Foundation. The research studies the effects of Kaizen events and what makes them successful. Your company is one of the few companies chosen for the research and will get first access to the results. Your company will use the results to design better Kaizen events and better support its Kaizen event teams.

This survey asks for your opinions on the on the Kaizen event you just completed. The survey should take about 15 minutes to complete.

Participation in this survey is voluntary. You are not asked to provide your name, just the name of your Kaizen event team. The survey confidential and the privacy of your answers will be protected. No one at your company will see your individual answers.

If you wish to participate, please answer the questions on the next two pages (page 2 and page 3). The questions ask you to describe how much you agree or disagree with statements about the goals of your team. If a question does not seem to be applicable, please answer "disagree" or "strongly disagree."

In the survey, "Kaizen event" refers to your Kaizen event team and not to any other teams working at the same time. "Work area" refers to the manufacturing area that is the focus of your Kaizen event.

Thank you for your help in this important research! If you have any questions or comments, please contact [insert name of University contact, e-mail, and phone number].

Kaizen Team Name ______________________________	Strongly Disagree 1	Disagree 2	Tend to Disagree 3	Tend to Agree 4	Agree 5	Strongly Agree 6
Circle the response that BEST describes your opinion.						
1. Most members of our [insert name of Kaizen event] team would like to be part of Kaizen events in the future.	1	2	3	4	5	6
2. Our [insert name of Kaizen event] team spent a lot of time discussing ideas before trying them out in the work area	1	2	3	4	5	6
3. Overall, this Kaizen event increased our [insert name of Kaizen event] team members' interest in our work.	1	2	3	4	5	6
4. Overall, this Kaizen event increased our [insert name of Kaizen event] team members' knowledge of the need for continuous improvement.	1	2	3	4	5	6
5. This Kaizen event has improved the performance of this work area.	1	2	3	4	5	6
6. Our [insert name of Kaizen event] team was free to make changes to the work area as soon as we thought of them.	1	2	3	4	5	6
7. Most of our [insert name of Kaizen event] team members gained new skills as a result of our participation in this Kaizen event.	1	2	3	4	5	6
8. Our [insert name of Kaizen event] team had enough help from our facilitator to get our work done.	1	2	3	4	5	6
9. Our [insert name of Kaizen event] team spent as much time as possible in the work area	1	2	3	4	5	6
10. Our [insert name of Kaizen event] team had a lot of freedom in determining how to improve this work area.	1	2	3	4	5	6
11. Most of our [insert name of Kaizen event] team members liked being part of this Kaizen event.	1	2	3	4	5	6
12. Our [insert name of Kaizen event] team had enough equipment to get our work done.	1	2	3	4	5	6
13. This Kaizen event increased most of our [insert name of Kaizen event] team members' ability to measure the impact of changes made to this work area.	1	2	3	4	5	6
14. Overall, this Kaizen event increased our [insert name of Kaizen event] team members' knowledge of what continuous improvement is.	1	2	3	4	5	6
15. Our [insert name of Kaizen event] team had a lot of freedom in determining how we spent our time during the event.	1	2	3	4	5	6
16. In general, this Kaizen event increased my team members' knowledge of our role in continuous improvement.	1	2	3	4	5	6
17. Most of our [insert name of Kaizen event] team members can communicate new ideas about improvements as a result of our participation in this Kaizen event.	1	2	3	4	5	6
18. In general, this Kaizen event increased our [insert name of Kaizen event] team members' knowledge of how continuous improvement can be applied.	1	2	3	4	5	6

	Strongly Disagree 1	Disagree 2	Tend to Disagree 3	Tend to Agree 4	Agree 5	Strongly Agree 6
Circle the response that BEST describes your opinion.						
19. Our [insert name of Kaizen event] team had a lot of freedom in determining what changes to make to this work area.	1	2	3	4	5	6
20. Our [insert name of Kaizen event] team tried out changes to the work area right after we thought of them	1	2	3	4	5	6
21. Overall, this Kaizen event helped people in this area work together to improve performance.	1	2	3	4	5	6
22. This work area improved measurably as a result of this Kaizen event.	1	2	3	4	5	6
23. Our [insert name of Kaizen event] team spent very little time in our meeting room	1	2	3	4	5	6
24. Our [insert name of Kaizen event] team had enough materials and supplies to get our work done	1	2	3	4	5	6
25. This Kaizen event made most of our [insert name of Kaizen event] team members more comfortable working with others to identify improvements.	1	2	3	4	5	6
26. Our [insert name of Kaizen event] team had enough help from others in our organization to get our work done.	1	2	3	4	5	6
27. In general, this Kaizen event motivated the members of our [insert name of Kaizen event] team to perform better.	1	2	3	4	5	6
28. This Kaizen event had a positive effect on this work area.	1	2	3	4	5	6
29. Our [insert name of Kaizen event] team had enough contact with management to get our work done.	1	2	3	4	5	6

31. How many Kaizen events total have you participated in? ***(Fill in the blank)*** ________________________________

32. Which functional area most closely describes your current job? ***(Circle one number)***
    1. OPERATIONS (ex. Production Associate, Maintenance Technician, etc.)
    2. NON-OPERATIONS (ex. Engineer, Marketing Associate, Manager, etc.)

33. What were the biggest obstacles to the success of the [insert name of Kaizen event] team (what did your team have to work hardest to overcome)?

34. What were the biggest contributors to the success of the [insert name of Kaizen event] team (what most helped your team to complete its work)?

Thank you for your participation! If you have any other comments about your experience with Kaizen events, please include them on the back of this page.

APPENDIX J: FINAL VERSION OF REPORT OUT SURVEY

Report Out Survey

Hello [insert name of Kaizen event] Team Member,

Your help is needed on this important research. This survey is part of a research project sponsored by the National Science Foundation. The research studies the effects of Kaizen events and what makes them successful. Your company is one of the few companies chosen for the research and will get first access to the results. Your company will use the results to design better Kaizen events and better support its Kaizen event teams.

This survey asks for your opinions on the Kaizen event you just completed (the [insert name of Kaizen event] event). The survey should take about 15 minutes to complete.

Please use the same survey code you used on the Kickoff Survey. Again, including this code is very important to allow the researchers to analyze survey results, but will not identify who you are, since only you know which code you chose. If you did not participate in the Kickoff Survey, please use one of the next three options to choose your survey code:

Option 1: The first four letters of your mother's maiden name [ex. Brown = "brow"]
Option 2: The month and day of your birthday [ex. January 1 = "0101"]
Option 3: The first four letters of your pet or child's name [ex. Dolly = "doll"]

Write your survey code on the top of the next page (page 2), after the words "Survey Code." **Do not include your name.** The survey is confidential and the privacy of your answers will be protected. No one at your company will see your individual answers.

Participation in this survey is voluntary. If you wish to participate, please answer the questions on the next three pages (pages 2 - 4) and return the survey in the envelope provided. The questions ask you to describe how much you agree or disagree with statements about the goals of your team. You may decline to answer any question(s) you choose. If a question does not seem to apply to your team, it may be that you "disagree" or "strongly disagree" with the statement.

In the survey, "Kaizen event" refers to your Kaizen event team and not to any other teams working at the same time. "Work area" refers to the manufacturing area that is the focus of your Kaizen event.

Thank you for your help in this important research! If you have any questions or comments, please contact [insert name of University contact, e-mail, and phone number]. If you have questions about your rights as a research participant, please contact the Virginia Tech Institutional Review Board (IRB) Human Protections Administrator, [insert name, e-mail and phone number of university Institutional Review Board (IRB) Human Protections Administrator].

Kaizen Team Name: [insert name of Kaizen event] Survey Code:	Strongly Disagree 1	Disagree 2	Tend to Disagree 3	Tend to Agree 4	Agree 5	Strongly Agree 6
Circle the response that BEST describes your opinion.						
1. Our [insert name of Kaizen event] team spent a lot of time discussing ideas before trying them out in the work area.	1	2	3	4	5	6
2. Most of our [insert name of Kaizen event] team members liked being part of this Kaizen event.	1	2	3	4	5	6
3. In general, our [insert name of Kaizen event] team members are comfortable working with others to identify improvements in this work area.	1	2	3	4	5	6
4. Our [insert name of Kaizen event] team valued the diversity in our team members.	1	2	3	4	5	6
5. Our [insert name of Kaizen event] team had enough help from others in our organization to get our work done.	1	2	3	4	5	6
6. Our [insert name of Kaizen event] team respected each others' feelings.	1	2	3	4	5	6
7. Most of our [insert name of Kaizen event] team members gained new skills as a result of participation in this Kaizen event.	1	2	3	4	5	6
8. Our [insert name of Kaizen event] team spent very little time in our meeting room.	1	2	3	4	5	6
9. Our [insert name of Kaizen event] team had enough equipment to get our work done.	1	2	3	4	5	6
10. In general, this Kaizen event increased our [insert name of Kaizen event] team members' knowledge of our role in continuous improvement.	1	2	3	4	5	6
11. In general, this Kaizen event motivated the members of our [insert name of Kaizen event] team to perform better.	1	2	3	4	5	6
12. Our [insert name of Kaizen event] team spent as much time as possible in the work area.	1	2	3	4	5	6
13. Our [insert name of Kaizen event] team had a lot of freedom in determining what changes to make to this work area.	1	2	3	4	5	6
14. Our [insert name of Kaizen event] team had enough materials and supplies to get our work done.	1	2	3	4	5	6
15. Most of our [insert name of Kaizen event] team members are able to measure the impact of changes made to this work area.	1	2	3	4	5	6
16. Overall, this Kaizen event increased our [insert name of Kaizen event] team members' interest in work.	1	2	3	4	5	6
17. Our [insert name of Kaizen event] team had enough contact with management to get our work done.	1	2	3	4	5	6
18. Our [insert name of Kaizen event] team valued each member's unique contributions.	1	2	3	4	5	6
19. This Kaizen event has improved the performance of this work area.	1	2	3	4	5	6

	Strongly Disagree 1	Disagree 2	Tend to Disagree 3	Tend to Agree 4	Agree 5	Strongly Agree 6
Circle the response that BEST describes your opinion.						
20. Our [insert name of Kaizen event] team respected each others' opinions.	1	2	3	4	5	6
21. Most members of our [insert name of Kaizen event] team would like to be part of Kaizen events in the future.	1	2	3	4	5	6
22. In general, this Kaizen event increased our [insert name of Kaizen event] team members' knowledge of how continuous improvement can be applied.	1	2	3	4	5	6
23. Most of our [insert name of Kaizen event] team members can communicate new ideas about improvements as a result of participation in this Kaizen event.	1	2	3	4	5	6
24. Overall, this Kaizen event increased our [insert name of Kaizen event] team members' knowledge of what continuous improvement is.	1	2	3	4	5	6
25. Our [insert name of Kaizen event] team had a lot of freedom in determining how we spent our time during the event.	1	2	3	4	5	6
26. Overall, this Kaizen event helped people in this area work together to improve performance.	1	2	3	4	5	6
27. This work area improved measurably as a result of this Kaizen event.	1	2	3	4	5	6
28. Our [insert name of Kaizen event] team had enough help from our facilitator to get our work done.	1	2	3	4	5	6
29. Our [insert name of Kaizen event] team was free to make changes to the work area as soon as we thought of them.	1	2	3	4	5	6
30. Our [insert name of Kaizen event] team communicated openly.	1	2	3	4	5	6
31. Our [insert name of Kaizen event] team tried out changes to the work area right after we thought of them.	1	2	3	4	5	6
32. Our [insert name of Kaizen event] team had a lot of freedom in determining how to improve this work area.	1	2	3	4	5	6
33. Overall, this Kaizen event increased our [insert name of Kaizen event] team members' knowledge of the need for continuous improvement.	1	2	3	4	5	6
34. This Kaizen event had a positive effect on this work area.	1	2	3	4	5	6
35. Overall, this Kaizen event was a success.	1	2	3	4	5	6

36. Including this event, how many Kaizen events total have you participated in? ***(Fill in the blank)***

____

37. Which functional area most closely describes your current job? ***(Circle one number)***

1. OPERATIONS (ex. Production Associate, Maintenance Technician, etc.)
2. NON-OPERATIONS (ex. Engineer, Marketing Associate, Manager, etc.)

38. What were the biggest obstacles to the success of the [insert name of Kaizen event] team (what did your team have to work hardest to overcome)?

39. What were the biggest contributors to the success of the [insert name of Kaizen event] team (what most helped your team to complete its work)?

Thank you for your participation! If you have any other comments about your experience with Kaizen events, please include them on the back of this page.

APPENDIX K: PILOT VERSION OF EVENT INFORMATION SHEET

Event Information Sheet

Hello Kaizen Event Facilitator,

Your help is needed on this important research. This questionnaire is part of a research project sponsored by the National Science Foundation. The research studies the effects of Kaizen events and what makes them successful. Your company is one of the few companies chosen for the research and will get first access to the results. You will be able to use the results to design better Kaizen events and better support Kaizen event teams.

This questionnaire asks for information about a specific Kaizen event that you facilitated. It should only take about 10 minutes to complete.

Participation in this research is voluntary. If you wish to participate, please answer the questions on the next three pages (pages 2 - 4).

Thank you for your help in this important research! If you have any questions or comments, please contact [insert name of University contact, e-mail, and phone number].

Event Information Sheet

1. Event Name: ____________________________

2. Dates: _______________________

3. Team Composition

Please fill-in the **number** of Kaizen event team members in each job category:

# Operators	# Technicians	# Engineers	# Managers	# Supervisors	# Other

4. Team Leader

How many Kaizen events **total** has the team leader led or co-led in the past three years? __________

5. Work Area Complexity

Please answer the following questions about the work area that was the target of this Kaizen event:

Circle the response that BEST describes your opinion.	Strongly Disagree	Disagree	Tend to Disagree	Tend to Agree	Agree	Strongly Agree
The work the target work area does is routine.	1	2	3	4	5	6
The target work area produces the same product (SKU) most of the time.	1	2	3	4	5	6
A given product (SKU) requires the same processing steps each time it is produced.	1	2	3	4	5	6
Most of the products (SKUs) produced in the work area follow a very similar production process.	1	2	3	4	5	6

6. Event Planning

How many hours **total** did you and others spend planning this Kaizen event? _____________

Were there any **unusual** planning activities completed for this event, which are not normally part of the event planning process? If so, what? (***Please briefly describe below***)

Example: "*pre-event meetings with work area employees"* (if this is not usually part of the event planning process)

7. Team Goals

Please list the goals of the Kaizen event team and the actual team results.

Please also indicate which goals describe the main purpose of the event (the "major" or "most important" goals) and which goals were less important (secondary goals).

Example

Team Goal	Result Achieved	Main Goal or Secondary Goal?
1. Reduce Cycle Time by 90%	70% reduction in cycle time	Main Goal
2. Achieve a 3S rating	Achieved a 3S rating	Secondary Goal

This Kaizen Event Team:

Team Goal	Result Achieved	Main Goal or Secondary Goal?
1.		
2.		
3.		
4.		
5.		
6.		
7.		
8.		
9.		
10.		

8. Event Success

Circle the response that BEST describes your opinion.	Strongly Disagree	Disagree	Tend to Disagree	Tend to Agree	Agree	Strongly Agree
Overall, this Kaizen event was a success.	1	2	3	4	5	6

9. Management Interaction with the Kaizen Event Team

Please briefly describe below the degree of face-to-face interaction members of management had with the Kaizen event team. (Ex. Did members of management attend the kickoff and report out meetings? Did members of management visit the Kaizen team while it was at work?)

10. Team Use of Problem Solving Tools

During the event, how many hours **total** did you spend with the team (**not** including the kickoff meeting, training and the report-out meeting)? _______________

Please list **all** problem solving tools used by the team in the box below. Then, ***for each tool***, please rate the team's use of the tool on: 1) ***appropriateness*** of using this tool to address the team's goals; and 2) ***quality*** of the team's use of this tool.

	How ***appropriate*** was this tool for the team's objectives?						How ***well*** did the team use this tool?					
	Completely Inappropriate	Inappropriate	Somewhat Inappropriate	Somewhat Appropriate	Appropriate	Completely Appropriate	Very Poorly	Poorly	Somewhat Poorly	Somewhat Well	Well	Very Well
Tool 1:	1	2	3	4	5	6	1	2	3	4	5	6
Tool 2:	1	2	3	4	5	6	1	2	3	4	5	6
Tool 3:	1	2	3	4	5	6	1	2	3	4	5	6
Tool 4:	1	2	3	4	5	6	1	2	3	4	5	6
Tool 5:	1	2	3	4	5	6	1	2	3	4	5	6
Tool 6:	1	2	3	4	5	6	1	2	3	4	5	6
Tool 7:	1	2	3	4	5	6	1	2	3	4	5	6
Tool 8:	1	2	3	4	5	6	1	2	3	4	5	6
Tool 9:	1	2	3	4	5	6	1	2	3	4	5	6
Tool 10:	1	2	3	4	5	6	1	2	3	4	5	6

APPENDIX L: FINAL VERSION OF EVENT INFORMATION SHEET

Event Information Sheet

Hello [insert name of Kaizen event] Event Facilitator,

Your help is needed on this important research. This questionnaire is part of a research project sponsored by the National Science Foundation. The research studies the effects of Kaizen events and what makes them successful. Your company is one of the few companies chosen for the research and will get first access to the results. You will be able to use the results to design better Kaizen events and better support Kaizen event teams.

This questionnaire asks for information about a specific Kaizen event that you facilitated (the [insert name of Kaizen event] event). It should only take about 15 minutes to complete.

Participation in this research is voluntary. If you wish to participate, please answer the questions on the next five pages (pages 2 – 6). You may decline to answer any question(s) you choose.

Thank you for your help in this important research! If you have any questions or comments, please contact [insert name of University contact, e-mail, and phone number].

If you have questions about your rights as a research participant, you should contact the Institutional Review Board (IRB) Human Protections Administrator at [insert name, e-mail and phone number of university Institutional Review Board (IRB) Human Protections Administrator].

Event Information Sheet

1. **Event Name:** [insert name of Kaizen event]

2. **Dates:**

3. **Team Composition**

Please fill-in the **number** of [insert name of Kaizen event] team members in each job category. If any team members fall in the "Other" category please briefly describe their job category (e.g., "Marketing") or role (e.g., "Facilitator").

# Operators	# Technicians	# Engineers	# Managers	# Supervisors	# Other

How many of the team members were <u>from the target work area</u>?

# Operators	# Technicians	# Engineers	# Managers	# Supervisors	# Other

How many of the team members were <u>new to (unfamiliar with) the target work area</u>?

# Operators	# Technicians	# Engineers	# Managers	# Supervisors	# Other

4. **Team Leader**

Including the [insert name of Kaizen event], how many Kaizen events **total** has the team leader led or co-led in the past three years?

5. **Work Area Complexity**

Please answer the following questions about the work area that was the target of the [insert name of Kaizen event] event:

Work Area Name:

Please provide a very brief description of the major products of the work area, the relative percentage of each produced (of the total work area volume) and the units of cycle time for each (seconds, minutes, hours, days, weeks, or months).

Product Description	% of Total Work Area Volume	Cycle Time Units
1.		
2.		
3.		
4.		
5.		
6.		
7.		
8.		
9.		
10.		

Please answer the questions below about the complexity of the target work area's production processes:

Circle the response that BEST describes your opinion.	Strongly Disagree	Disagree	Tend to Disagree	Tend to Agree	Agree	Strongly Agree
The work the target work area does is routine.	1☐	2☐	3☐	4☐	5☐	6☐
The target work area produces the same product (SKU) most of the time.	1☐	2☐	3☐	4☐	5☐	6☐
A given product (SKU) requires the same processing steps each time it is produced.	1☐	2☐	3☐	4☐	5☐	6☐
Most of the products (SKUs) produced in the work area follow a very similar production process.	1☐	2☐	3☐	4☐	5☐	6☐

Have there been other Kaizen events that have targeted this work area (before the [insert name of Kaizen event])? 1 ☐ = YES 2 ☐ =NO

If yes, please indicate the general description and date (MM/YY) of the other events. Also, if you are familiar with the event, please indicate your opinion of its general level of success. If you are not familiar with the event, please select "Don't Know."

Kaizen Event Description	**Date (MM/YY)**	**Very Unsuccessful**	**Unsuccessful**	**Somewhat Unsuccessful**	**Somewhat Successful**	**Successful**	**Very Successful**	**Don't Know**
Ex. Lathe Setup Reduction	*09/04*	1☐	2☐	3☐	4☐	5☒	6☐	☐
1.		1☐	2☐	3☐	4☐	5☐	6☐	☐
2.		1☐	2☐	3☐	4☐	5☐	6☐	☐
3.		1☐	2☐	3☐	4☐	5☐	6☐	☐
4.		1☐	2☐	3☐	4☐	5☐	6☐	☐
5.		1☐	2☐	3☐	4☐	5☐	6☐	☐
6.		1☐	2☐	3☐	4☐	5☐	6☐	☐
7.		1☐	2☐	3☐	4☐	5☐	6☐	☐
8.		1☐	2☐	3☐	4☐	5☐	6☐	☐

6. Team Goals

Please list the goals of the [insert name of Kaizen event] team and the actual team results.

Please also indicate which goals describe the main purpose of the [insert name of Kaizen event] event (the "major" or "most important" goals) and which goals were less important (secondary goals).

Example

Team Goal	**Result Achieved**	**Main Goal or Secondary Goal?**
1. Reduce Cycle Time by 90%	70% reduction in cycle time	Main Goal
2. Achieve a 3S rating	Achieved a 3S rating	Secondary Goal

The [insert name of Kaizen event] Team:

Team Goal	Result Achieved	Main Goal or Secondary Goal?
1.		
2.		
3.		
4.		
5.		
6.		
7.		
8.		
9.		
10.		

7. Event Success

Circle the response that BEST describes your opinion.	Strongly Disagree	Disagree	Tend to Disagree	Tend to Agree	Agree	Strongly Agree
Overall, the [insert name of Kaizen event] event was a success.	1☐	2☐	3☐	4☐	5☐	6☐

8. Team Use of Problem Solving Tools

Did the [insert name of Kaizen event] team follow a structured improvement methodology (e.g., SMED, Standard Work, etc.)? 1 ☐ = YES 2 ☐ =NO
If so, which methodology?
During the [insert name of Kaizen event] event, how many hours total did you spend with the team (not including the kickoff meeting, training and the report-out meeting)?

Please list all problem solving tools used by the [insert name of Kaizen event] team in the box below (e.g., brainstorming, fishbone diagramming, Pareto analysis, SMED, etc.). Then, for each tool, please rate the team's use of the tool on: 1) appropriateness of using this tool to address the team's goals; and 2) quality of the team's use of this tool.

	How ***appropriate*** was this tool for the team's objectives?						How do you rate the overall ***quality*** of the team's use of this tool?					
	Very Inappropriate	Inappropriate	Somewhat Inappropriate	Somewhat Appropriate	Appropriate	Very Appropriate	Very Poor	Poor	Fair	Good	Very Good	Excellent
Tool 1:	1☐	2☐	3☐	4☐	5☐	6☐	1☐	2☐	3☐	4☐	5☐	6☐
Tool 2:	1☐	2☐	3☐	4☐	5☐	6☐	1☐	2☐	3☐	4☐	5☐	6☐
Tool 3:	1☐	2☐	3☐	4☐	5☐	6☐	1☐	2☐	3☐	4☐	5☐	6☐
Tool 4:	1☐	2☐	3☐	4☐	5☐	6☐	1☐	2☐	3☐	4☐	5☐	6☐
Tool 5:	1☐	2☐	3☐	4☐	5☐	6☐	1☐	2☐	3☐	4☐	5☐	6☐
Tool 6:	1☐	2☐	3☐	4☐	5☐	6☐	1☐	2☐	3☐	4☐	5☐	6☐
Tool 7:	1☐	2☐	3☐	4☐	5☐	6☐	1☐	2☐	3☐	4☐	5☐	6☐
Tool 8:	1☐	2☐	3☐	4☐	5☐	6☐	1☐	2☐	3☐	4☐	5☐	6☐
Tool 9:	1☐	2☐	3☐	4☐	5☐	6☐	1☐	2☐	3☐	4☐	5☐	6☐
Tool 10:	1☐	2☐	3☐	4☐	5☐	6☐	1☐	2☐	3☐	4☐	5☐	6☐

9. Team Decision Making Process

Please indicate your perceptions of the overall quality of the [insert name of Kaizen event] team's decision making process.

Circle the response that BEST describes your opinion.	Strongly Disagree	Disagree	Tend to Disagree	Tend to Agree	Agree	Strongly Agree
The process the [insert name of Kaizen event] team used to make decisions was sound.	1☐	2☐	3☐	4☐	5☐	6☐

10. Event Pre-Work

How many hours **total** did you and others spend planning and doing other pre-work for the [insert name of Kaizen event] event? (collecting current state data, developing event plan, etc.)

Were there any **unusual** pre-work activities completed for the [insert name of Kaizen event] event, which are not normally part of the event pre-work process? If so, what? (**Please briefly describe below**)

Example: *"pre-event meetings with work area employees*" (if this is not usually part of the event planning process)

11. Kickoff Meeting Process

Circle the response that BEST describes your opinion.	Not at All					To a Great Extent
To what extent did the [insert name of Kaizen event] team interact during the kickoff meeting? (i.e., to what extent was there group discussion of the event goals).	1☐	2☐	3☐	4☐	5☐	6☐

Please **briefly** describe the kickoff meeting agenda below.

12. Team Training

What topics were covered in team training? (e.g., group rules/norms, SMED process, seven wastes of lean, etc.)

Did everyone on the team receive the same training? 1☐ = YES 2☐ =NO
If no, please briefly explain below.

13. Management Interaction with the Kaizen Event Team

Please briefly describe below the degree of face-to-face interaction members of management had with the [insert name of Kaizen event] team. (Did members of management attend the kickoff and report out meetings? Did members of management visit the Kaizen team while it was at work?)

14. Biggest Obstacles to Team Success

In your opinion, what were the biggest <u>obstacles</u> to the success of the [insert name of Kaizen event] team? (That is, what factors – either internal or external to the team – most strongly threatened the success of the team?)

15. Biggest Contributors to Team Success

What were the biggest <u>contributors</u> to the success of the [insert name of Kaizen event] team? (That is, what factors – either internal or external to the team – most increased the likelihood of team success?)

APPENDIX M: PILOT VERSION OF KAIZEN EVENT PROGRAM INTERVIEW GUIDE

Kaizen Event Program Interview Guide

Interviewer:
Date:
Start time:
End time:

Organizational Demographics:
Name:
Location:
Industry sector:
Major products:
Number of employees:
Number of local facilities:

Introductory Comments

This interview is part of a research project sponsored by the National Science Foundation, in which your company his participating. The purpose of the research is to study the effects of Kaizen events and what makes them successful. Your company is one of the few companies chosen for the research and will get first access to the results. Your company will be able to use the research results to design better Kaizen events and better support its Kaizen event teams.

The **purpose** of this interview is to **obtain information about your company's overall experience with Kaizen events to date**. Interview questions will cover a number of different aspects of Kaizen events – planning, the event process, sustainability mechanisms, etc. This information will help researchers understand the context and process of Kaizen events within your company. This will help the researchers better interpret research findings and understand the similarities and differences between participating companies.

This interview should take about 30- 40 minutes. Your participation is voluntary and confidential. You can decide not to participate at any time. In addition, the privacy of your answers will be protected. No one at the organization will see your answers to these interview questions.

Do you wish to continue?

Background Questions

1. When was the first kaizen event in your company?

2. How long has the company been sponsoring kaizen events systematically?

3. How many kaizen events have you had in the company?

4. What types of objective benefits/results have you realized from Kaizen events?

5. What types of non-measurable benefits have you realized?

6. To what extent are kaizen events viewed as a success in your organization?

Event Planning

7. In general, how often are Kaizen events conducted in your company?

8. Does your company keep a master schedule of all events planned for some upcoming time window?
    a. If yes, <u>ask for a copy of the schedule</u>
    b. If yes, what is the time window of the event schedule

9. What have typically been the catalysts for change determining the need for kaizen events? (customer-driven? Competition-driven? Etc.)

10. How do you select which work areas should be targeted for Kaizen events? (e.g., what is your "filtering"/Kaizen event selection process?)

11. How do you know when not to do a Kaizen event (boundary conditions)?

12. What percent of your organization has experienced kaizen events? (% by production line/cell?)

    a. What is the relative percentage of events in non-manufacturing versus manufacturing areas?

    b. What are the major types of processes (e.g., operations, sales, engineering) in which kaizen events have been conducted? What is the relative percentage of events that have been conducted in each?

13. What percent have had two or more events?

14. What percent of your workforce has been involved in kaizen events?

Event Process

15. What types of events does your company conduct (e.g., SMED, Std Wrk)

16. If your company conducts different types of events, what is the relative percentage of each type?

17. Is there a separate, standardized process for each type of event (e.g. separate training material for SMED versus Std Wrk), or all are all events conducted using the same general process?

18. For each type of Kaizen event, what is the total time frame (in days)?

19. Does your company conduct shorter (one or two day) versions of Kaizen events that are not considered "formal events"?

    a. If yes, what boundary conditions are used to differentiate "formal" versus "informal" events?

20. What is the typical format for each type of kaizen event and how long is spent on each part of the event? (e.g., kick-off? Training? Analysis? Designing future state?)

21. Do you have a formal report out process?

    a. If yes, ask for an example report out file (ideally, one for several different types, types of areas – manufacturing versus non-manufacturing)

22. What types of performance areas (measures) are typically targeted in each type of kaizen event?

23. For each type of event, how many people typically make up the kaizen team?

24. For each type of event, what is the typical composition of the kaizen team?

25. What are the typical roles within a Kaizen event team?

    a. Do all events use the same person as the facilitator, or does this role rotate within the organization or even external to the organization?

26. What mechanisms do you have in place to sustain Kaizen event outcomes?

27. Have there been any issues (problems, difficulties) with sustainability (either of measurable or non-measurable benefits)?

28. Other problem areas with events?

29. How did you develop your event process (external consultant, training from parent company, published literature, etc.)?

30. How have you improved your Kaizen process over the years?

31. What are the resources available for kaizen events? (Budget? Facilitation? Training?)

32. What % of your budget and/or what % of people's time is devoted to Kaizen events on an ongoing basis?

33. What is your process for capturing "lessons learned"?
    a. For a single Kaizen event
    b. Across multiple Kaizen events

34. Do you see your use of Kaizen events increasing, decreasing or staying the same over the next several years?

35. Polar opposite question: what characterizes your best versus "worst" Kaizen events?

APPENDIX N: PILOT VERSION OF KAIZEN EVENT PROGRAM INTERVIEW GUIDE – WRITTEN STATEMENT FOR PARTICIPANTS

Interview for Kaizen Event Research:
Written Statement for Participants

Hello,

This interview is part of a research project sponsored by the National Science Foundation, in which your company is participating. The purpose of the research is to study the effects of Kaizen events and what makes them successful. Your company is one of the few companies chosen for the research and will get first access to the results. Your company will be able to use the research results to design better Kaizen events and better support its Kaizen event teams.

The **purpose** of this interview is to **obtain information about your company's overall experience with Kaizen events to date**. Interview questions will cover a number of different aspects of Kaizen events – planning, the event process, sustainability mechanisms, etc. This information will help researchers understand the context and process of Kaizen events within your company. This will help the researchers better interpret research findings and understand the similarities and differences between participating companies.

This interview should take about 30 - 40 minutes. Your participation is voluntary and confidential. You can decide not to participate at any time. In addition, the privacy of your answers will be protected. No one at the organization will see your answers to these interview questions.

To participate, please contact [insert name of University contact, e-mail, and phone number]. Thank you for your help in this important research! If you have any questions or comments, please contact [insert name of University] using the information above.

APPENDIX O: FINAL VERSION OF KAIZEN EVENT PROGRAM INTERVIEW GUIDE

Kaizen Event Program Interview Guide

Interviewer:
Date:
Start time:

End time:

Organizational Demographics:
Name:
Location:
Industry sector:
Major products:
Number of employees:
Number of local facilities:

Introductory Comments
This interview is part of a research project sponsored by the National Science Foundation, in which your company is participating. The purpose of the research is to study the effects of Kaizen events and what makes them successful. Your company is one of the few companies chosen for the research and will get first access to the results. Your company will be able to use the research results to design better Kaizen events and better support its Kaizen event teams.

The **purpose** of this interview is to **obtain information about your company's overall experience with Kaizen events to date**. Interview questions will cover a number of different aspects of Kaizen events – planning, the event process, sustainability mechanisms, etc. This information will help researchers understand the context and process of Kaizen events within your company. This will help the researchers better interpret research findings and understand the similarities and differences between participating companies.

Do you wish to continue?

Background Questions

1. When was the first Kaizen event in your company?

2. How long has the company been sponsoring kaizen events systematically?

3. How many Kaizen events have you had in the company?

4. What types of objective benefits/results have you realized from Kaizen events?

5. What types of non-measurable benefits have you realized?

6. To what extent are Kaizen events viewed as a success in your organization?

Event Planning

7. In general, how often are Kaizen events conducted in your company?

8. Does your company keep a master schedule of all events planned for some upcoming time window?
    a. If yes, <u>ask for a copy of the schedule</u>
    b. If yes, what is the time window of the event schedule

9. What have typically been the catalysts for change determining the need for Kaizen events? (customer-driven? Competition-driven? Etc.)

10. How do you select which work areas should be targeted for Kaizen events? (e.g., what is your "filtering"/Kaizen event selection process?)

11. How do you know when not to do a Kaizen event (boundary conditions)?

12. What percent of your organization has experienced Kaizen events? (% by production line/cell?)

    a. What is the relative percentage of events in non-manufacturing versus manufacturing areas?

    b. What are the major types of processes (e.g., operations, sales, engineering) in which kaizen events have been conducted? What is the relative percentage of events that have been conducted in each?

13. What percent have had two or more events?

14. What percent of your workforce has been involved in Kaizen events?

Event Process

15. What types of events does your company conduct (e.g., SMED, Std Wrk)?

16. If your company conducts different types of events, what is the relative percentage of each type?

17. Is there a separate, standardized process for each type of event (e.g. separate training material for SMED versus Std Wrk), or are all events conducted using the same general process?

18. For each type of Kaizen event, what is the total time frame (in days)?

19. Does your company conduct shorter (one or two day) versions of Kaizen events that are not considered "formal events"?

    a. If yes, what boundary conditions are used to differentiate "formal" versus "informal" events?

20. What is the typical format for each type of Kaizen event and how long is spent on each part of the event? (e.g., kick-off? Training? Analysis? Designing future state?)

21. Do you have a formal report out process?

    a. If yes, ask for an example report out file (ideally, one for several different types, types of areas – manufacturing versus non-manufacturing)

22. What types of performance areas (measures) are typically targeted in each type of Kaizen event?

23. For each type of event, how many people typically make up the Kaizen team?

24. For each type of event, what is the typical composition of the Kaizen team?

25. What are the typical roles within a Kaizen event team?

    a. Do all events use the same person as the facilitator, or does this role rotate within the organization or even external to the organization?

26. What mechanisms do you have in place to sustain Kaizen event outcomes?

27. Have there been any issues (problems, difficulties) with sustainability (either of measurable or non-measurable benefits)?

28. Other problem areas with events?

29. How did you develop your event process (external consultant, training from parent company, published literature, etc.)?

30. How have you improved your Kaizen process over the years?

31. What are the resources available for kaizen events? (Budget? Facilitation? Training?)

32. What % of your budget and/or what % of people's time is devoted to Kaizen events on an ongoing basis?

33. What is your process for capturing "lessons learned"?
    a. For a single Kaizen event
    b. Across multiple Kaizen events

34. Do you see your use of Kaizen events increasing, decreasing or staying the same over the next several years?

35. Polar opposite question: what characterizes your best versus "worst" Kaizen events?

APPENDIX P: FINAL VERSION OF KAIZEN EVENT PROGRAM INTERVIEW GUIDE – WRITTEN STATEMENT FOR PARTICIPANTS

Kaizen Event Program Interview: Informed Consent Form

Hello,

This interview is part of a research project sponsored by the National Science Foundation, in which your company is participating. This form describes what is involved in this interview, so you can make an informed choice of whether or not to participate. You will be provided with a copy of this form for your records.

I. Purpose

The purpose of the research is to study the effects of Kaizen events and what makes them successful. The purpose of this interview is to obtain information about your company's overall experience with Kaizen events to date. This information will help the researchers better interpret research findings and understand the similarities and differences between participating companies. The research project will consist of 7 – 15 companies in a variety of industries. To participate, the companies must meet the following selection criteria: 1) manufacture products of some type; 2) have conducted Kaizen events for at least one year; 3) use Kaizen events systematically, as part of a formal organizational strategy; and 4) conduct at least one Kaizen event a month. All companies who meet the criteria are eligible to participate. The identity of participating companies will be kept confidential. In any publications, participating companies will be referred to by a code name (ex. "Company A"), which only the researchers will know.

II. Procedure

This interview is a one-time event and it should take about 30 - 40 minutes. The interview will occur in your office, either in person or over the phone. Interview questions will cover a number of different aspects of Kaizen events – planning, the event process, sustainability mechanisms, etc. For the questions that you chose to answer, you are asked to answer based on your knowledge of your company's Kaizen event program. With your permission, the research team will create an audiotape recording of the interview. The audiotape will only be used to enable the researchers to accurately transcribe the interview. The audiotape will not be released to anyone other than the researchers and will be destroyed once the interview has been transcribed (within one week of the interview).

III. Risks and Benefits.

The risks from participation in this study are no greater than those encountered in daily life. There are several ways your company can benefit from your participation. Your company is one of the few companies participating in this research project and will get first access to the results. Your company will be able to use the research project results to design better Kaizen events and better support its Kaizen event teams. This description of the risks and benefits has been provided to help you more fully understand what is involved in this interview and in the research project -- it is not offered as compensation (incentive) for your participation.

IV. Confidentiality of Participation

The confidentiality of your participation this study will be preserved. No one at your company will be informed of your decision of whether or not to participate. The audiotape will be stored in a secured location at [insert university name] under the supervision of [insert PI name]. The audiotape will be translated into notes by a member of the research team and then destroyed, within one week of the interview. The notes will contain only the organization name and employee position. Any information published from the interview will contain only a code name for the organization (ex. "Company A"), which only the research team will know; any other identifying information (product names, company terms, etc.) will also be removed.

V. Compensation

No compensation (either financial or non-financial) will be provided for participation.

VI. Freedom to Withdraw

You may choose not to participate or choose not to answer any question at any time without penalty.

I understand that my participation in this study is completely voluntary and that I may choose not to participate or choose not to answer any question at any time without penalty. I voluntarily agree to participate in this study. My only responsibility in this study is to answer all the questions that I chose to answer based on my knowledge of my company's Kaizen event program.

I understand that any questions I have about the research study or specific procedures should be directed to [insert name of university contacts, addresses, e-mails and phone number – investigator, PI, & departmental reviewer].

If I have questions about my rights as a research participant, I should contact the Institutional Review Board (IRB) Human Protections Administrator at [insert e-mail and phone number of university Institutional Review Board].

My signature below indicates that I have read and that I understand the process described above and give my informed and voluntary consent to participate in this study. I understand that I will receive a signed copy of this consent form.

Signature of Participant ____

Printed Name of Participant ____

Date Signed ____

This Informed Consent is valid from ________ to ________.

APPENDIX Q: ADMINISTRATION AND TRAINING TOOLS FOR ORGANIZATIONAL FACILITATORS

Kickoff Survey Administration Script

Dear Facilitator:

Please read the following statement aloud to the Kaizen event team when you hand out the Kickoff Survey.

"This survey is part of a research project at Virginia Tech and Oregon State University that studies the effects of Kaizen events and what makes them successful. [Insert company name] is one of the few companies chosen for the research. [Insert company name] thinks this research is important, and we'll use the results to improve our Kaizen events and how we support Kaizen event teams.

This survey asks for *your* opinions on the goals of your Kaizen event team.

Please take a couple minutes to read over the instructions on the first page of the survey. [*Wait until people have finished reading*].

The survey asks you to choose a unique survey code that will allow the researchers to perform analysis on the survey results. This code is very important. It will <u>not</u> be used to identify you, since only you will know what your code is. If you have already used a survey code on previous surveys for this research, please use <u>the same</u> code. If you have not completed any previous surveys for this research, please use <u>one</u> of the following three options to choose your survey code:

Option 1: The first four letters of your mother's maiden name

- [*ex. Brown = "brow"*]

Option 2: The month and day of your birthday

- [*ex. January 1 = "0101"*]

Option 3: The first four letters of your pet or child's name

- [*ex. Dolly = "doll"*]

Please remember the code you select, since it will be used for future surveys. **Do not include your name.** The survey is <u>confidential</u> and the privacy of your answers will be protected. No one at [insert company name] will see your individual answers.

Does anyone have any other questions about the survey instructions? [*Answer questions then finish reading script*].

I will leave the room to give you time to complete the survey.

Participation in this survey is voluntary. You can choose to stop participating at any time. After I leave, if you wish to participate, please complete the 'Kickoff Survey' and return it to this envelope. [*Hold up envelope*]. You may decline to answer any question(s) you wish.

Thank you!

Report-Out Survey Administration Script

Dear Facilitator:

Please read the following statement aloud to the Kaizen event team when you hand out the Report-Out Survey.

> "Like the Kickoff Survey, this survey is part of the research project at Virginia Tech and Oregon State University that studies Kaizen events. The process for completing the survey is the same as the process for completing the Kickoff Survey. Again, [Insert company name] thinks this research is important, and we'll use the results to improve our Kaizen events and how we support Kaizen event teams.
>
> This survey asks for *your* opinions about the Kaizen event you just completed.
>
> Like last time, please take a couple minutes to read over the instructions on the first page of the survey. [*Wait until people have finished reading*].
>
> Please use <u>the same</u> survey code you used on the Kickoff Survey. Again, including this code is very important to allow the researchers to analyze survey results, but will <u>not</u> identify who you are, since only you know which code you chose. If you did not participate in the Kickoff Survey, please use <u>one</u> of the following three options to choose your survey code:
>
> Option 1: The first four letters of your mother's maiden name
> - [*ex. Brown = "brow"*]
>
> Option 2: The month and day of your birthday
> - [*ex. January 1 = "0101"*]
>
> Option 3: The first four letters of your pet or child's name
> - [*ex. Dolly = "doll"*]
>
> **Do not include your name.** The survey is <u>confidential</u> and the privacy of your answers will be protected. No one at [insert company name] will see your individual answers.
>
> Does anyone have any other questions about the survey instructions? [*Answer questions then finish reading script*].
>
> I will leave the room to give you time to complete the survey.
>
> Again, participation in this survey is voluntary. You can choose to stop participating at any time. After I leave, if you wish to participate, please complete the 'Report-Out Survey' and return it to this envelope. [*Hold up envelope*]. You may decline to answer any question(s) you choose.
>
> Thank you!

APPENDIX R: TABLE OF EVENTS STUDIED BY COMPANY

Code: TPM = Total Productive Maintenance; PI = (General) Process Improvement; SMED = Setup Reduction; VSM = Value Stream Mapping; 5S = Housekeeping/Work Area Organization; L = Layout

Company A

Event	Dates	Method(s)	Target System	Focus	Team Size	Response Kickoff	Response Report Out
1. TPM 1A (501)	10/23/05 – 10/28/05	TPM	One Machine	Improving the condition of the target machine and training operators in TPM	5	5 (100%)	4 (80%)
2. PI 1A (502)	10/31/05 – 11/04/05	Standard Work	Manufacturing Process	Identifying root causes of scrap and implementing countermeasures	6	6 (100%)	5 (83%)
3. PI 2A (504)	12/05/05 – 12/08/05	Process Flow, SMED	Manufacturing Process	Improving the material flow of small lot sizes through a bottleneck process	11	10 (91%)	8 (73%)
4. TPM 2A (505)	12/12/05 – 12/1605	TPM	One Machine	Improving the condition of the target machine and training operators in TPM	4	2 (50%)	4 (100%)
5. TPM 3A (506)	01/08/06 – 01/1106	TPM	Three Machines	Improving the condition of the target machines and training operators in TPM	7	7 (100%)	6 (86%)
6. SMED 1A (509)	01/16/06 – 01/20/06	SMED	One Machine	Reducing changeover times for machine	8	8 (100%)	6 (75%)
7. SMED 2A (510)	01/23/06 – 01/27/06	SMED, 5S	One Machine	Reducing changeover times for machine	5	5 (100%)	4 (80%)
8. PI 3A (512)	02/06/06 – 02/14/06	None	Manufacturing Process	Improving material flow through the manufacturing process	10	9 (90%)	8 (80%)
9. TPM 4A (514)	03/13/06 – 03/17/06	TPM	Two Machines	Improving the condition of the target machines and training operators in TPM	7	6 (86%)	6 (86%)
10. PI 4A (517)	03/21/06 – 03/23/06	None/ Brainstorming	Manufacturing Process/ Department	Creating a future state layout for target department and developing an implementation plan	4	4 (100%)	4 (100%)
11. PI 5A (520)	03/27/06 – 03/30/06	Six Sigma	Manufacturing Process	Creating a future state layout for target	6	6 (100%)	6 (100%)

				process and developing an implementation plan			
12. TPM 5A (521)	04/24/06 – 04/28/06	TPM	Two Machines	Improving the condition of the target machines and training operators in TPM	5	5 (100%)	4 (80%)
13. PI 6A (523)	05/01/06 – 05/05/06	None	Manufacturing Process/ Department	Prepare a designated location for two new pieces of machinery (location decided in advance of event)	6	6 (100%)	4 (67%)
14. PI 7A (532)	05/15/06 – 05/18/06	SMED	Family of machines	Establishing standard setup and inspection procedures for target machines, including developing training aids for setups	5	4 (80%)	3 (60%)
15. PI 8A (530)	05/22/06 – 05/26/06	None	Manufacturing Process/Depart ment	Laying out a cell for a new product line and installing equipment in cell	6	4 (67%)	4 (67%)

Company B

Event	Dates	Method(s)	Target System	Focus	Team Size	Response Kickoff	Response Report Out
1. PI 1B (319)	03/28/06 – 03/30/06	TPI	Service process	Improving and standardizing the inquiry to quote process for standard product lines	12	9 (75%)	11 (92%)
2. VSM 1B (322)	04/03/06 - 04/05/06	VSM	Manufacturing process (product repair)	Create a current state map of target process (repair process) and identify general areas/triggers for improvement	11	11 (100%)	9 (82%)
3. PI 2B (324)	05/08/06 – 05/12/06	Standard Work	Manufacturing process	Reducing cell lead-time	22	14 (64%)	11 (50%)
4. PI 3B (325)	05/08/06 – 05/12/06	Standard Work	Manufacturing process	Redesigning cell layout to meet a specified takt rate	8	6 (75%)	6 (75%)
5. PI 4B (326)	05/08/06 – 05/12/06	Standard Work	Manufacturing process	Redesigning cell layout to meet a specified takt rate	10	9 (90%)	6 (60%)
6. PI 5B (327)	05/08/06 – 05/12/06	Standard Work	Manufacturing process	Simplifying the process and reducing changeover times	11	9 (73%)	6 (55%)
7. PI 6B (328)	05/08/06 – 05/12/06	Standard Work	Manufacturing process	Improving material and information flow within the cell	20	19 (95%)	16 (80%)
8. PI 7B (329)	05/08/06 – 05/12/06	Standard Work	Service process	Reducing time for fax filing and distribution	7	7 (100%)	6 (86%)

Company C

Event	Dates	Method(s)	Target System	Focus	Team Size	Response Kickoff	Response Report Out
1. PI C1 (615)	01/09/06 – 01/13/06	None/JIT and Kaizen	Manufacturing process common across multiple product lines	Balancing the line using a certain bottleneck machine as pacesetter for takt time	11	8 (73%)	9 (82%)
2. PI C2 (634)	1/23/06 - 1/27/06	None	Manufacturing process	Increasing process flexibility and reducing batching by replacing two dedicated machines with a more flexible model	8	8 (100%)	7 (88%)
3. PI C3 (635)	2/13/06 - 2/20/06	Standard Work	Manufacturing process	Determining job standards for cell and training operators to meet standards (task improvement and allocation)	8	8 (100%)	3 (38%)
4. PI C4 (618)	2/27/06 - 3/3/06	Standard Work	Manufacturing process	Determining job standards for cell and training operators to meet standards (task improvement and allocation)	9	9 (100%)	5 (56%)
5. PI C5 (616)	3/13/06 - 3/17/06	Standard Work	Manufacturing process	Determining job standards for cell and training operators to meet standards (task improvement and allocation)	7	7 (100%)	7 (100%)
6. PI C6 (641)	3/27/06 - 3/31/06	None	Manufacturing process	Reducing lead-time variance across different products	5	5 (100%)	3 (60%)
7. PI C7 (636)	4/10/06 - 4/14/06	Standard Work and DMAIC	Manufacturing process common across multiple product lines	Reducing defects	4	0 (0%[13])	0 (0%)
8. PI C8 (637)	4/24/06 - 4/28/06	None/ continuous flow	Manufacturing process	Reducing cycle time	5	5 (100%)	2 (40%)
9. 5S 1C (638)	5/8/206 - 5/12/06	5S	Manufacturing process/depart ment	Improving inventory management and reducing defects	4	2 (50%)	3 (60%)
10. PI C9 (639)	5/22/06 - 5/26/06	None/7 wastes and waste reduction, lean line design and kanban systems	Manufacturing process	Improving process flow and area staffing requirements	5	5 (100%)	4 (80%)

[13] A response rate of zero percent indicates that none of these surveys were returned by organization

11. 5S 2C (640)	6/12/06 - 6/16/06	5S	Manufacturing process/department	Improving inventory management and reducing part retrieval time	5	5 (100%)	0 (0%)

Company D

Event	Dates	Method(s)	Target System	Focus	Team Size	Response Kickoff	Response Report Out
1. PI 1D (811)	01/23/06 - 01/26/06	DMAIC, Lean	Manufacturing process	Improving the condition of the target machine and training operators in TPM	13	11 (85%)	11 (85%)
2. PI 2D (813)	02/20/06 – 02/24/06	Lean	Manufacturing /supply process	Identifying root causes of scrap and implementing countermeasures	11	11 (100%)	10 (91%)
3. PI 3D (831)	4/18/06 – 4/20/06	VSM	One tool within an engineering/ quality assurance process	Improving the material flow of small lot sizes through a bottleneck process	10	10 (100%)	6 (60%)
4. PI 4D (833)	06/12/06 – 06/14/06	Lean	Engineering/ supply process (part specification and purchasing)	Improving the condition of the target machine and training operators in TPM	8	8 (100%)	6 (88%)

Company E

Event	Dates	Method(s)	Target System	Focus	Team Size	Response Kickoff	Response Report Out
1. PI 1E (100)	12/20/05 – 12/22/05	Cellular Design	Manufacturing process/department	Redesigning cell layout to improve product flow and reduce cycle time	4	3 (75%)	3 (75%)
2. PI 2E (104)	03/14/06 – 03/17/06	Standard Work and Processing Mapping	Service process	Reducing cycle time of customer quote process	8	7 (88%)	4 (50%)
3. VSM 1E (101)	3/14/06 – 3/16/06	VSM	Service process	Documenting the current state, designing a future state and identifying future Kaizen events to implement future state	4	4 (100%)	4 (100%)
4. SMED 1E (102)	03/20/06 – 03/22/06	SMED	One machine	Reducing changeover time for target machine	5	5 (100%)	5 (100%)
5. PI 3E (106)	03/28/06 - 03/31/06	Standard Work	Manufacturing support process (ordering)	Improving part ordering process to reduce part shortages	5	5 (100%)	5 (100%)
6. PI 4E	04/03/06 –	None	Manufacturing	Reducing lead-time	4	4 (100%)	4 (100%)

(103)	04/05/06		process	and implementing one piece flow			
7. VSM 2E (107)	04/17/06 - 04/19/06	VSM	Manufacturing /shipping process	Improving shipping ("kitting") process to eliminate omissions in orders	5	4 (80%)	4 (80%)
8. PI 5E (105)	04/24/06 - 04/28/06	Standard Work	Manufacturing process	Reducing defects	8	7 (88%)	6 (75%)
9. PI 6E (108)	05/16/06 – 05/18/06	Standard Work	Manufacturing process	Reducing cycle time	6	5 (83%)	5 (83%)
10. PI 7E (109)	06/13/06 – 06/15/06	Process Mapping, Flow	Manufacturing support process (ordering)	Reducing cycle time of ordering process	7	6 (86%)	5 (71%)
11. SMED 2E (111)	6/22/06 – 6/23/06	SMED	One machine	Reducing changeover time for target machine	5	5 (100%)	5 (100%)
12. P1 8E (110)	6/26/06 – 6/28/06	Standard Work	Service Process	Reduce process complexity (number of steps) and cycle time for the target process	4	4 (100%)	3 (75%)

Company F

Event	Dates	Method(s)	Target System	Focus	Team Size	Response Kickoff	Response Report Out
1. L 1F (400)	01/11/06 – 01/12/06	None	Inventory Storage Area	Redesigning the layout of a storage area	7	7 (100%)	7 (100%)
2. TPM 1F (401)	03/24/06 – 03/25/06	TPM	One Machine	Developing an autonomous maintenance program for the target machine	6	5 (83%)	5 (83%)
3. 5S 1F (402)	03/28/06 – 03/30/06	6S	Manufacturing Process/Department	Raising 5S (6S) score of target cell	4	4 (100%)	4 (100%)
4. 5S 2F (403)	04/17/06 – 04/18/06	6S	Manufacturing Process/Department	Implementing 5S (6S) to improve organization of target cell	3	3 (100%)	3 (100%)
5. PI 1F (404)	05/10/06 – 05/12/06	None	Manufacturing Process	Documenting current state (operation times, etc.) and improving material flow through the target process	8	8 (100%)	8 (100%)
6. TPM 2F (405)	05/12/06 – 05/13/06	TPM	One Machine	Developing an autonomous maintenance program for the target machine	6	6 (100%)	6 (100%)

APPENDIX S: SUMMARY OF STUDY VARIABLE RESULTS BY COMPANY

Org		AT	Task KSA	Overall Success	% Goals Met	IMA	GC	GDF	TA	Functional Het.	Team Mbr KE	Team Lder Exp	MS	Event Plan Proc	Work Area Routineness	ACC	AO	IP	Tool Approp.	Tool Quality	% Goals Met_o	Team Mbr KE_o	Team Lder Exp_o	Event Plan Proc_o
A	mean	5.20	5.16	4.93	0.00	5.11	4.90	3.68	4.91	0.53	0.32	2.13	5.02	0.88	4.83	5.06	4.79	5.35	5.27	4.79	0.99	4.08	3.13	11.27
	median	5.17	5.29	5.00	0.00	5.17	4.76	3.70	5.08	0.60	0.30	1.00	5.11	0.78	4.75	5.08	4.88	5.45	5.00	5.00	1.00	4.00	2.00	6.00
	max	5.83	5.53	6.00	0.00	5.56	5.56	4.19	5.33	0.99	1.15	13.00	5.50	1.78	6.00	5.61	5.67	5.73	6.00	6.00	1.00	9.75	14.00	60.00
	min	4.58	4.65	1.00	-0.04	3.83	4.45	2.67	3.42	0.05	0.00	0.00	4.28	0.30	4.50	4.47	3.63	4.80	4.50	4.00	0.91	1.11	1.00	2.00
	std dev	0.35	0.30	1.28	0.01	0.41	0.34	0.41	0.50	0.28	0.37	3.78	0.45	0.36	0.37	0.33	0.52	0.28	0.43	0.54	0.03	2.43	3.78	14.45
B	mean	4.92	4.84	5.38	-0.10	4.78	4.51	4.22	4.88	0.69	0.49	3.13	4.57	1.33	3.91	4.82	4.15	5.12	5.49	4.77	0.88	5.32	4.13	32.25
	median	4.93	4.78	5.00	0.00	4.76	4.65	4.22	4.95	0.73	0.54	2.50	4.57	1.19	3.75	4.86	4.50	5.25	5.54	5.00	1.00	5.40	3.50	15.50
	max	5.33	5.40	6.00	0.00	5.47	4.89	4.42	5.40	0.92	1.04	10.00	5.00	2.08	6.00	5.17	5.06	5.49	6.00	5.67	1.00	8.33	11.00	120.00
	min	4.40	4.27	5.00	-0.70	3.60	3.93	3.93	3.72	0.23	0.00	0.00	3.98	1.08	1.75	4.46	2.36	4.51	4.80	3.80	0.20	1.71	1.00	12.00
	std dev	0.37	0.40	0.52	0.24	0.58	0.35	0.17	0.53	0.20	0.37	3.31	0.31	0.37	1.76	0.26	0.97	0.33	0.51	0.75	0.28	1.89	3.31	38.31
C	mean	5.00	4.89	5.00	-0.09	4.75	4.16	3.75	4.52	0.36	0.07	0.29	4.72	0.38	4.14	4.56	4.29	4.92	5.29	4.32	0.86	2.50	1.29	2.79
	median	5.17	4.70	5.00	0.00	4.83	4.23	3.58	4.42	0.35	0.00	0.00	4.52	0.48	4.50	4.75	4.38	5.00	5.00	4.00	1.00	2.22	1.00	3.00
	max	5.67	5.41	6.00	0.00	5.56	4.65	4.53	5.22	0.69	0.48	2.00	5.33	0.60	4.75	4.97	4.83	5.60	6.00	5.00	1.00	4.88	3.00	4.00
	min	4.00	4.44	4.00	-0.30	3.48	3.44	3.38	4.00	0.00	0.00	0.00	4.17	-0.30	1.50	3.61	3.71	3.69	5.00	4.00	0.50	1.00	1.00	0.50
	std dev	0.63	0.35	1.00	0.15	0.69	0.41	0.40	0.39	0.20	0.18	0.76	0.48	0.31	1.17	0.50	0.40	0.63	0.39	0.47	0.24	1.18	0.76	1.15
D	mean	4.76	4.81	4.50	0.00	4.70	4.55	3.69	4.81	0.49	0.15	0.50	4.93	0.73	3.19	4.70	3.08	5.15	5.44	4.88	1.00	3.13	1.50	5.75
	median	4.89	4.88	5.50	0.00	4.79	4.63	3.67	4.97	0.45	0.15	0.50	5.06	0.65	3.13	4.62	3.04	5.24	5.38	5.00	1.00	2.84	1.50	4.50
	max	5.07	5.11	6.00	0.00	5.20	4.90	3.98	5.19	0.61	0.30	1.00	5.28	1.00	3.50	5.09	3.70	5.47	6.00	6.00	1.00	4.10	2.00	10.00
	min	4.20	4.39	1.00	0.00	4.04	4.04	3.43	4.12	0.44	0.00	0.00	4.30	0.60	3.00	4.45	2.55	4.64	5.00	3.50	1.00	2.75	1.00	4.00
	std dev	0.41	0.35	2.38	0.00	0.56	0.36	0.27	0.50	0.08	0.17	0.58	0.45	0.19	0.24	0.29	0.50	0.36	0.43	1.31	0.00	0.65	0.58	2.87
E	mean	4.84	4.52	4.45	-0.33	4.80	4.54	3.31	4.79	1.12	0.85	13.09	4.64	0.75	4.39	4.81	3.95	5.12	5.69	4.29	0.66	17.16	14.09	9.18
	median	4.89	4.53	5.00	-0.09	4.73	4.56	2.83	4.87	1.06	1.00	9.00	4.67	0.60	4.50	4.77	4.13	5.10	5.80	4.50	0.82	11.40	10.00	4.00
	max	5.42	5.19	6.00	0.00	5.67	4.85	4.81	5.50	1.54	1.71	50.00	5.28	1.60	6.00	5.33	5.00	5.40	6.00	5.50	1.00	34.50	51.00	40.00
	min	4.33	4.06	2.00	-2.00	4.20	4.25	2.51	4.33	0.50	0.00	0.00	4.00	0.30	2.50	4.33	2.33	4.55	5.00	3.00	0.00	3.17	1.00	2.00
	std dev	0.38	0.35	1.04	0.57	0.50	0.19	0.78	0.34	0.36	0.61	15.31	0.39	0.42	1.28	0.34	0.81	0.26	0.40	0.64	0.31	11.50	15.31	11.50
F	mean	5.08	4.86	5.17	-0.02	5.06	4.58	3.16	4.92	0.60	0.44	3.00	5.17	0.83	5.04	4.89	4.59	5.16	5.39	4.78	0.96	4.68	4.00	17.00
	median	5.13	4.83	5.00	0.00	4.84	4.52	3.39	4.77	0.64	0.45	2.00	5.13	0.60	5.38	4.81	4.38	5.23	5.17	5.00	1.00	4.42	3.00	4.00
	max	5.67	5.63	6.00	0.00	5.78	5.25	4.22	5.56	0.95	1.04	10.00	5.58	1.90	6.00	5.67	5.38	5.80	6.00	5.33	1.00	9.00	11.00	80.00
	min	4.33	3.94	4.00	-0.11	4.50	4.00	1.92	4.40	0.10	0.00	0.00	4.96	0.48	3.25	4.25	4.00	4.46	5.00	4.00	0.78	1.25	1.00	3.00
	std dev	0.56	0.62	0.75	0.04	0.53	0.44	0.89	0.43	0.29	0.41	3.79	0.22	0.55	1.04	0.48	0.57	0.45	0.49	0.62	0.09	2.59	3.79	30.92
OVERALL	mean	5.00	4.87	4.90	-0.10	4.91	4.59	3.63	4.82	0.47	0.67	0.43	4.84	0.83	4.39	4.85	4.28	5.17	5.42	4.62	0.88	6.88	5.37	13.19
	median	5.08	4.80	5.00	0.00	5.00	4.63	3.73	4.87	0.45	0.66	0.30	4.90	0.78	4.50	4.83	4.40	5.25	5.33	4.60	1.00	4.54	2.00	6.00
	max	5.83	5.63	6.00	0.00	5.78	5.56	4.81	5.56	0.86	1.54	1.71	5.58	2.08	6.00	5.67	5.67	5.80	6.00	6.00	1.00	34.50	51.00	120.00
	min	4.00	3.94	1.00	-2.00	3.48	3.44	1.92	3.42	0.13	0.00	0.00	3.98	-0.30	1.50	3.61	2.33	3.69	4.50	3.00	0.00	1.00	1.00	0.50
ANOVA p		0.30	0.01	0.59	0.11	0.44	0.00	0.01	0.52	0.00	0.00	0.00	0.03	0.00	0.06	0.09	0.00	0.25	0.26	0.31	n/a	n/a	n/a	n/a
Kruskal-Wallis p		0.26	0.01	0.43	0.00	0.41	0.01	0.01	0.27	0.01	0.00	0.03	0.04	0.00	0.06	0.21	0.00	0.35	0.30	0.37				

- AT = Attitude
- Task KSA = Task Knowledge Skills and Attitudes
- Overall Success = Overall Perceived Success, facilitator rating of overall event success
- % of Goals Met = percentage of major improvement goals (*log transformed*)
- IMA = Impact on Area
- GC = Goal Clarity
- GDF = Goal Difficulty
- TA = Team Autonomy
- Functional Het. = Functional Heterogeneity; index from 0 – 1 measuring cross functional diversity of team
- Team Mbr KE = Team Kaizen Experience; average number of total events participated in per team member (including current event) (*log transformed*)
- Team Lder Exp = Team Leader Experience; total number of events led (including current event) (*log transformed*)
- MS = Management Support
- Event Plan Proc = Event Planning Process, total hours spent planning the event (*log transformed*)
- Work Area Routineness = facilitator rating of the predictability of the target work area in four different dimensions
- ACC = Affective Commitment to Change
- AO = Action Orientation
- IP = Internal Processes
- Tool Approp. = Tool Appropriateness, average facilitator rating for appropriateness of the problem-solving tools used by the team
- Tool Quality = facilitator rating for quality of the problem-solving tools used by the team

% Goals Met_o, Team Mrb KE_o, Team Ldr Exp_o, and Event Plan Proc_o are the original (untransformed) values of these variables

The following variables are measured on a scale of 1 = “strong disagree” to 6 = “strongly agree”: AT, Task KSA, Overall Success, IMA, GC, GDF, TA, MS, Work Area Routineness, ACC, AO, IP, Tool Approp., and Tool Quality

Full scale values:
1 = “strong disagree”
2 = “disagree”
3 = “tend to disagree”
4 = “tend to agree”
5 = “agree”
6 = “strongly agree”

APPENDIX T: FULL CORRELATION ANALYSIS RESULTS

Response (Predictor)	**Correlation Coefficient**	**P-value for $\widehat{\beta}_{GEE}$**	**Intraclass Correlation Coefficient (ρ)**	**Correlation Coefficient_OLS**
Attitude (Task KSA)	.710	.0000	-.069	.712*
Task KSA (Attitude)	.711	.0000	.215	.712*
Overall Perceived Success (Attitude)	.122	.3023	-.044	.125
Attitude (Overall Perceived Success)	.123	.4068	.037	.125
Overall Perceived Success (Task KSA)	.166	.1193	-.054	.173
Task KSA (Overall Perceived Success)	.155	.3326	.280	.173
Impact on Area (Attitude)	.632	.0000	-.071	.643*
Attitude (Impact on Area)	.639	.0000	-.052	.643*
Impact on Area (Task KSA)	.690	.0000	.014	.690*
Task KSA (Impact on Area)	.688	.0000	.299	.690*
Impact on Area (Overall Perceived Success)	.224	.1003	.009	.224
Overall Perceived Success (Impact on Area)	.224	.1073	-.036	.224
% of Goals Met (Attitude)	.000	.8569	.107	.035
Attitude (% of Goals Met)	.000	.9937	.044	.035
% of Goals Met (Task KSA)	.000	.3451	.172	.030
Task KSA (% of Goals Met)	.000	.1879	.363	.030
% of Goals Met (Overall Perceived Success)	.170	.2982	.084	.172
Overall Perceived Success (% of Goals Met)	.163	.0767	-.061	.172
% of Goals Met (Impact on Area)	-.019	.4377	.123	-.055
Impact on Area (% of Goals Met)	-.048	.5882	.027	-.055

* = significant at the $\alpha = 0.05/10 = 0.005$ level using OLS estimates and n = 51