AN ABSTRACT OF THE THESIS OF

Rebecca M. Mitchell for the degree of Honors Baccalaureate of Science in Industrial Engineering and Manufacturing Engineering presented on March 9, 2007. Title: A Validity Study for the Kaizen Event Kick-Off Survey.

Abstract Approved:

____

Toni L. Doolen

This study was conducted to support a research project sponsored by the National Science Foundation (NSF). The purpose of the NSF research project was to determine how to design kaizen events that result in both short term and long term performance improvements. One of the tools developed to evaluate the effectiveness of different kaizen events was a kick-off survey. The use of surveys has been criticized in the past because many researchers do not verify that surveys are valid. The purpose of this investigation was to evaluate the validity of the kaizen event kick-off survey.

The validity of the kaizen event kick-off survey was evaluated with a validity survey. The validity survey was given to university students who then rated how well each of the kick-off survey items measured the underlying kick-off survey constructs. The ratings from the validity survey were analyzed using factor analysis and Wilcoxon signed ranks test. The factor analysis indicated that all of the items except for GDF7, ACC3, and ACC5 demonstrated content validity and construct validity. The Wilcoxon signed ranks test demonstrated content validity for all of the items.

Key Words: validity, construct, survey, lean manufacturing, kaizen event

Corresponding e-mail address: mitchere@onid.orst.edu

©Copyright by Rebecca M. Mitchell
March 9, 2007
All Rights Reserved

A Validity Study for the Kaizen Event Kick-Off Survey

by

Rebecca M. Mitchell

A PROJECT

submitted to

Oregon State University

University Honors College

in partial fulfillment of
the requirements for the
degree of

Honors Baccalaureate of Science in
Industrial Engineering and Manufacturing Engineering
(Honors Associate)

Presented March 9, 2007
Commencement June 2007

Honors Baccalaureate of Science in Industrial Engineering and Manufacturing Engineering project of Rebecca M. Mitchell presented on March 9, 2007.

APPROVED:

Mentor, representing Industrial and Manufacturing Engineering

---

Committee Member, representing Industrial and Manufacturing Engineering

---

Committee Member, representing Industrial and Manufacturing Engineering

---

Interim Head, Department of Industrial and Manufacturing Engineering

---

Dean, University Honors College

I understand that my project will become part of the permanent collection of Oregon State University, University Honors College. My signature below authorizes release of my project to any reader upon request.

Rebecca M. Mitchell, Author

ACKNOWLEDGEMENTS

I would like to thank my mentor, Dr. Toni L. Doolen for her guidance and support. I would also like to thank my committee members, Dr. Sundar Atre and June Worley.

TABLE OF CONTENTS

Page

1. Introduction.. 1

 1.1 Background and Significance.. 1
 1.2 Research Topic and Findings.. 2

2. Literature Review... 3

 2.1 Lean Manufacturing... 3
 2.2 Kaizen... 5
 2.3 Surveys as Research Instruments.................................... 8
 2.4 Validity and Reliability... 9

3. Methodology... 15

 3.1 Research Instrument.. 15
 3.2 IRB Approval... 17
 3.3 Data Collection Details.. 17
 3.4 Participants... 18
 3.5 Data Collection and Analysis....................................... 18

4. Research Findings... 21

 4.1 Demographic Research Findings...................................... 21
 4.2 Factor Analysis.. 21
 4.3 Wilcoxon Signed Ranks Test... 30
 4.4 Recommendations.. 32
 4.5 Conclusion... 33

Bibliography... 34

Appendices... 36

 Appendix A.. 37
 Appendix B.. 43

LIST OF FIGURES

Figure		Page
4.2.1	Plot of rotated components in space of only two components shows that goal clarity items load separately than all the other items............................	25
4.2.2	Plot of rotated components in space of only two components shows that items for goal difficulty load separately than all the other items..................	27
4.2.3	Plot of rotated components in space of only two components shows that items for the affective commitment to change construct load differently than all the other items...	29

LIST OF TABLES

Table		Page
3.1.1	Constructs and Definitions	15
3.5.1	Codes of Item Statements	19
4.2.1	Rotated component matrix for item ratings when compared to the definition for goal clarity	23
4.2.2	Rotated component matrix for item ratings when compared to the definition for goal difficulty	26
4.2.3	Rotated component matrix for item ratings when compared to the definition for affective commitment to change	28
4.3.1	Results of the Wilcoxon signed ranks test	31

A Validity Study for the Kaizen Event Kick-Off Survey

1. INTRODUCTION

1.1 Background and Significance

Kaizen events have become increasingly popular in recent years as a way for manufacturing organizations to pursue improvements in their production systems. During these events, employees use lean manufacturing techniques to focus on eliminating waste in the manufacturing process. Kaizen events promote teamwork and empower employees to make changes in their organization. They have been shown to be an effective mechanism for improving the performance of manufacturing companies.

Dr. T. Doolen, a professor in the Department of Industrial and Manufacturing Engineering at Oregon State University, is working on a research project sponsored by the National Science Foundation (NSF) to determine how to design kaizen events that result in both short term and long term performance improvements. One of the tools developed to evaluate the effectiveness of different kaizen events is a kick-off survey. Employees at various manufacturing organizations complete the survey at the beginning of a kaizen event. A collaborative research team at Oregon State University (OSU) and Virginia Tech (VT) then uses this survey to help determine how employee perceptions at the start of a kaizen event are related to both improvements in the manufacturing process and to employee knowledge, skills and/or attitudes.

1.2 Research Topic and Findings

One of the concerns of the research team is whether or not the kick-off survey instrument is valid. Written survey items can be interpreted in many ways, and the research team must determine if the surveys are truly measuring the variables they were designed to measure. This thesis will describe the results of a research study designed to assess the validity of the kaizen event kick-off survey. In particular, as a result of this research, a study was conducted to test the relationship between the survey constructs and the written survey items.

Based on the completed study, evidence has been collected to support the validity of the kaizen event kick-off survey. The survey items were found to correlate with the proposed research constructs. In the next section, previous research related to the current study is reviewed. This is followed by a chapter detailing the study conducted. In the final section, the research findings from the study are presented and discussed. The thesis concludes with a summary of the significance of this research.

2. LITERATURE REVIEW

2.1 Lean Manufacturing

This study focused on determining the validity of a survey instrument used to measure the impact of kaizen events on social system outcomes. To understand the survey items, it is important to understand kaizen events. In addition, since kaizen events are used to help organizations become lean, it is also necessary to understand lean manufacturing.

Lean manufacturing is a term used to describe a manufacturing system that is primarily based on the Toyota Production System (Black & Hunter, 2003). The Toyota Production System (TPS) is a business strategy that was developed by Taiichi Ohno in the early 60's (Black & Hunter, 2003). The goal of this strategy was to reduce inventory and improve the quality and cost of Toyota's automobiles. As a result of this unique production system, Toyota became a leading automobile manufacturer (Black & Hunter, 2003). Other companies took notice, and the TPS became one of the most benchmarked production systems in the world (Black & Hunter, 2003).

Many companies in the United States have tried to create production systems that are largely based on the concepts and ideas brought about by Toyota. A few examples include the Ford Production System, the Chrysler Operation System, and GROWTTH which is the lean production system for a German company called Freudenberg-NOK (Shook, 1998; Day, 1998). These are all modern production systems based on TPS.

All of these companies have benchmarked TPS because it has been very successful, and it is drastically different from traditional mass manufacturing methods.

Lean manufacturing focuses on the customer rather than on the product. The entire purpose of lean manufacturing is to eliminate muda, the Japanese term for waste (Womack & Jones, 1996). Waste is considered to be any activity that absorbs resources without adding value to the product (Womack & Jones, 1996). Value is defined by the customer and therefore the goal is to eliminate all processes or activities which do not add value in the eyes of the customer.

The basic concept of lean is simple, but sometimes it is hard to grasp exactly how to use this concept to improve a production system. One way to think about lean manufacturing is to break it up into different lean principles. Lean Thinking, a popular book about lean manufacturing, breaks up lean into five main principles: value, the value stream, flow, pull, and perfection (Womack & Jones, 1996).

The first two principles of lean manufacturing focus on the value of a product from the eyes of the customer. Many companies focus on their own "operational 'efficiency'" and not on delivering a sound product to the customer (Womack & Jones, 1996, p. 34). Lean manufacturing stresses the importance of looking at a product from the customer's point of view. Then, from there, one can determine what parts of the manufacturing process are directly adding value to the product and which parts are not. The manufacturing processes that do not add value to the product are then eliminated.

The next key principle in lean manufacturing is to ensure that all of the remaining steps in the process flow together (Womack & Jones, 1996). This principle can be interpreted as a recommendation to arrange the manufacturing steps in such a way that makes the most sense. The key to flow is to focus on the product and not on the methods for making the product. It is more efficient and accurate to work on a product “continuously from raw material to finished good” than to handle the product multiple times using batch flow (Womack & Jones, 1996, p. 22).

The fourth lean principle introduced in Lean Thinking is pull. The idea behind pull is relatively simple. In order to avoid overproduction, do not make a product until a customer orders it. Let the customer *pull* the product from the manufacturer when they want it.

The final principle of lean production is perfection (Womack & Jones, 1996). Perfection simply means that the process is never entirely complete and that it is essential that companies continuously work to improve the production process. The ultimate goal is to make the production process perfect. However, due to changing customer needs, the production process never will be perfect, and a company must continuously improve the production process to ensure that it is always as close to perfect as possible.

2.2 Kaizen

A company can move closer to perfection by continuously working on improving its current production system. The part of lean manufacturing that focuses on continuous improvement is referred to as kaizen. Kaizen is the Japanese idea that “great improvement eventually comes from a series of small incremental gains” (Nicholas, 1998, p.37).

This is the purpose of holding a kaizen event. The goal is to improve the production process by holding workshops where employees “try to accomplish as much actual kaizen as possible” (Shook, 1998, p. 65). These events have become increasingly popular in recent years and are used by numerous companies to help implement and sustain lean manufacturing principles and practices. Some companies even try to measure success in an organization by counting the number of kaizen events held each year (Shook, 1998).

A prime example of a company which used kaizen events to implement lean manufacturing is Freudenberg-NOK. Freudenberg-NOK is the “largest manufacturer of sealing components in the world” (Day, 1998, p. 179). They created a lean production system in 1992, and four years later the company experienced a growth in sales of $400 million (Day, 1998). The CEO of the company attributed the sales growth to improved quality, cost and delivery which were attained by holding “2,500 kaizen events in 15 manufacturing plants” (Day, 1998, p. 179).

Other companies however, have had less success with kaizen events. Shook (1998) attributes this lack of success to the way that companies use them. Some companies hold kaizen events with the hope that the event will somehow “create a lean production system” (Rother, 1998, p. 492). However, a kaizen event by itself does not make a company lean.

Shook (1998) emphasizes some of the pitfalls of kaizen events in his article in the book Becoming Lean. He explains that companies need to use kaizen events as part of the plan for the entire system and not a group of stand-alone activities (Shook, 1998). Kaizen events can actually cause problems in a company if they are not part of an overall system (Shook, 1998). At the end of a kaizen event, there are usually “dozens of niggling problems that don’t surface until after the change” (Rother, 1998, p. 492). Manufacturing companies need to ensure that these problems are dealt with after a kaizen event to ensure that the event will have a positive effect on the company and its performance.

One of the keys to having successful kaizen events is to set up a kaizen program. Some suggestions for having a successful kaizen program are to have a lean champion and a governing committee (Ortiz, 2006). Ortiz (2006) suggest hiring an industrial or manufacturing engineer with expertise in lean manufacturing whose sole job is to execute kaizen. This person leads the kaizen events, is in charge of communicating with others in the organization about the events, and is responsible for following up on the events. Another suggestion by Ortiz is to use an event tracking worksheet to ensure that employees complete any necessary follow-up activities after the kaizen event (Ortiz, 2006).

A governing committee with managers from each of the business areas is a second recommendation towards an effective kaizen event program (Ortiz, 2006). The committee would be responsible for scheduling and supporting the kaizen events (Ortiz, 2006). It is very critical for managers to be involved in kaizen events in order to make them effective (Day, 1998; Ortiz, 2006).

A kaizen event can also be viewed as a way to train employees in lean manufacturing and to empower them to create positive changes in their own work areas. At Toyota, kaizen events were initially used more as a training tool to instill lean thinking than to actually make company improvements (Shook, 1998). It is important to train employees in lean thinking. Employee involvement and empowerment have been shown to be critical in creating a lean production facility (Liker, 1998).

2.3 Surveys as Research Instruments

Surveys are used in research to evaluate psychological behaviors and feelings. In the kaizen event study conducted by OSU and VT, surveys are used to evaluate the way that the kaizen event participants feel about a kaizen event. Surveys are a popular research tool, but many researchers fail to verify that they are both reliable and valid measures of the concepts which they are designed to measure (McGrath, 2005). It is impossible to accurately assess results if the tool used for measuring these results is faulty (McGrath, 2005). For this reason, it is important to determine whether or not a survey is valid.

Before discussing validity in detail, some of the common terminology associated with survey development and administration will be discussed. A survey instrument (a.k.a. questionnaire or survey) is composed of a series of questions or statements which are commonly referred to as items. These questions or statements are related to a theoretical concept of human behavior that the researcher is trying to evaluate (Babbie, 1998). This concept is referred to as the survey construct (Babbie, 1998).

If the item is a statement, then usually it is rated on a scale. Hinkin (1998) indicates that the most popular scale used for rating survey items is the Likert scale. Likert scales usually have five, seven or nine different ratings, but the five point scale is recommended because it has the highest reliablility. The scales have an odd number of ratings in order for the survey participant to select extreme values or a neutral number which is given by the middle value.

The people that take the surveys are referred to as survey participants or respondents. The person who administers the survey is referred to as the survey administrator.

2.4 Validity and Reliability

In order to draw appropriate conclusions from empirical studies that use surveys, it is important for the research team to demonstrate that a survey is both reliable and valid. Without this corroboration, an entire study can be rendered meaningless (Babbie, 1998). A survey is considered to be reliable if the results of the survey are consistent. In particular, if a survey is reliable, the results would be similar regardless of when and where the survey was administered (Muchinsky, 1997). Validity, conversely, is a measure of "accuracy and precision" (Babbie, 1998, p.94).

2.4.1 Reliability

Reliability is determined by evaluating how well the results of one item correlate to the results of other related items within a construct. One of the most common methods used to determine a survey's reliability is to calculate Cronbach's Alpha (Chow & Lui, 2003). An alpha value that is greater than 0.7 indicates that the survey is internally reliable (Chow & Lui, 2003).

2.4.2 Validity

Determining if a survey is valid is considerably more difficult. This may be due to the fact that there really is not a standard procedure for establishing validity (Hinkin, 1998). There are also many different types of validity and not all types of validity are necessarily relevant to every study.

Some of the most common types of validity include face validity, content validity, criterion-related validity, and construct validity (Muchinsky, 1997; Babbie, 1998). Face validity is the most basic type of validity. This type of validity is based on "common agreements and our individual mental images concerning a particular concept" (Muchinsky, 1997).

Another type of validity that is similar to face validity is referred to as content validity. Content validity is used to describe how well a measure accurately represents the domain of a construct (Babbie, 1998; Hinkin, 1998; Hinkin & Tracey, 1999). Although this seems straightforward, there are differing views on how to evaluate and measure content validity. This may be due to the fact that content validity is defined in different ways by different researchers.

There are two common views regarding content validity. The first one is that content validity means that the entire range of a construct domain has been sampled (Muchinsky, 1997). Determining if an entire domain has been sampled cannot be evaluated statistically. Thus, content validity, defined in this way, must be analyzed by somebody with knowledge of the domain (Chow & Lui, 2003). Yusof and Aspinwall (2000) believe that content validity should always be "subjectively evaluated by the researcher" (p. 453). Babbie (1998) makes a similar suggestion by saying that content validity can be established by "subject matter experts" (p. 96). This is consistent with Yusof and Aspinwalls' (2000) definition of content validity since the researcher is an expert on the subject and can therefore qualitatively evaluate the content validity by themselves.

A second view of content validity is that it is really just a part of construct validity. (Construct validity will be discussed in more detail later.) In this interpretation, content validity is associated with determining what content to include in a study, and in determining whether or not this content is reflective of the constructs. Hinkin and Tracey (1999) suggest that content validity is technically impossible to evaluate and that the researcher should focus on evaluating how well a "measure's items reflect a particular theoretical domain" (p.175). This type of content validity is also referred to as content adequacy and substantive validity (Hinkin & Tracey, 1999). These types of validity are all very similar to construct validity and are thus evaluated as a part of construct validity.

A third type of validity is referred to as criterion-related validity. It is "the extent to which a measuring instrument is related to an independent measure of relevant criteria" (Yusof & Aspinwall, 2000). It is also "called predictive validity" because this type of validity is ascertained when criteria can accurately be predicted based upon the measuring device (Muchinsky, 1997, p.133). Criterion-related validity can be difficult to establish when a comparative data set does not already exist (Babbie, 1998).

The final and most complex type of validity is construct validity (Babbie, 1998). The term, construct validity, was developed to help researchers evaluate the "representational accuracy of [survey instruments]" (McGrath, 2005, p. 113). It is defined as the “extent to which a [survey] measures what it is purported to measure” (Hinkin, 1998, p. 105).

Multiple sources agree that there are three steps to establishing construct validity (Serafini & Adams, 2002; Hinkin, 1998). The first is to clearly define the domain of a theoretical survey construct and develop survey items that represent the domain (Serafini & Adams, 2002; Hinkin, 1998). The second step is to assess the structure of the survey items and “empirically determine the extent to which these items measure” a survey construct (Serafini & Adams, 2002; Hinkin, 1998, pg.105). The final step is to examine whether the survey is externally valid by determining how well it predicts the variables it was intended to predict (Serafini & Adams, 2002; Hinkin 1998).

According to McGrath (2005), an important part of the first step to ensuring that a survey is valid is to carefully define a survey construct. A construct is based on the theoretical concept that is being evaluated, and can range from a very basic model of human behavior to a specific representation of a response to certain stimuli. How specific a construct is depends on what the researcher is evaluating, but it is important to ensure that the construct is precise because specific constructs are more valid than broad constructs. This is due to the fact that specific constructs are more easily represented by a succinct number of items. This also helps to ensure that the survey has content validity.

There are many different ways to go about creating survey constructs and items. Hinkin (1998) identified two common methods: the deductive approach and the inductive approach. In the deductive approach, a researcher thoroughly evaluates the subject of interest and then uses this evaluation to come up with a construct definition. Then, the researcher develops all the items needed to accurately measure this description. In doing this, the researcher must be careful to choose items that are measuring the entire domain. Again, this helps to establish content validity. In contrast, if an inductive approach is used, the researcher asks a group of respondents to discuss their feelings on a subject. Then, the researchers evaluate these feelings and attempt to categorize them. Each of the categories is then identified as a survey construct. Of the two methods, the deductive approach is considered to be the most valid because the method used to develop the items ensures that the items are consistent with the construct definition.

The type and number of items used is also important in establishing validity of the survey. Hinkin (1998) recommends using between four and six items for each construct because research indicates that using more items does not improve the validity of the construct and using fewer items can reduce the reliability of the results. Some researchers believe that as few as three items for each construct is sufficient (Serafini & Adams, 2002).

Guidelines also exist for developing items that ensure that the respondent can adequately rate them. Hinkin (1998) suggests making items simple and easy to understand. Each of the items used should also be redundant in the sense that each item should measure the same construct, but the survey creator should be wary of creating redundant items that vary in their level of difficulty (McGrath, 2005). Items that are hard for a respondent to evaluate may not be as reliable. The items should also be carefully developed so that each item relates to and measures only one construct. Additionally, the items should be worded neutrally, i.e. items should not lead the respondent (Hinkin, 1998).

Once a survey construct and relevant items have been created, the second step to establishing construct validity is to determine how well each of the items measures the construct it was designed to measure (Muchinsky, 1997; Babbie, 1998; Hinkin, 1998). Many different methods exist for doing this. Some of the methods used include exploratory factor analysis, confirmatory factor analysis, and analysis of variance. The details of some of these methods will be explained in greater detail later in the chapter on research findings.

The final step in establishing construct validity is to ensure that a survey produces predictable results. This can be done by determining how well items “correlate with other measures designed to assess similar constructs (convergent validity), and [how well] they do not correlate with dissimilar measures (discriminant validity)” (Hinkin, 1998, p.116). The most common method to evaluate discriminant and convergent validity is to use Multitrait-Multimethod Matrix (MTMM) (Hinkin, 1998). Many other methods, including factor analysis and analysis of variance, have also been used.

Once the three steps have been completed, the measuring instrument is considered to be valid (Hinkin, 1998). It should be remembered, however, that no single test of validity is completely effective. Surveys are based on psychological constructs that are concepts that only exist because they are agreed upon by a group of people (Muchinsky, 1997). Evaluating how well items measure corresponding constructs is therefore never entirely accurate because qualitative concepts are not necessarily “directly relatable to a quantitative evaluation” (McGrath, 2005, p.114). Therefore, validity tests cannot be viewed as necessarily proving or disproving the validity of a measure. Rather, these tests should be used to establish a certain degree of validity (Babbie, 1998).

3. METHODOLOGY

3.1 Research Instrument

A five page validity survey was used to evaluate correlations between the survey items and the survey constructs. The first page of the validity survey requested help from the participants to conduct the research and emphasized the fact that the survey was entirely voluntary. The first page also provided detailed instructions explaining how to complete the validity survey. A copy of the actual survey developed for this study is provided in Appendix A.

The survey was developed following the recommendations provided by Hinkin and Tracey (1999). The focus of the validity survey was to measure the content validity of the kick-off survey items. Each of the three pages of the validity survey had a survey construct along with its corresponding definition listed at the top of the page. The three survey constructs that were tested for this study include Goal Clarity, Goal Difficulty and Affective Commitment to Change. The constructs and their definitions are listed in Table 3.1.1. The construct definitions were developed by the collaborative research team as a part of a larger NSF study.

Table 3.1.1: Constructs and Definitions

Survey Construct	**Construct Definition**
Goal Clarity	These items describe team perceptions of the extent to which the kaizen event team's improvement goals have been clearly defined.
Goal Difficulty	These items describe team perceptions of the difficulty of the improvement goals set for the Kaizen event team.
Affective Commitment to Change (Commitment to Event Goals)	These items describe team perceptions of the need for the specific changes targeted by the kaizen event.

The survey items were listed below the survey constructs. Each respondent surveyed was asked to rate on a Likert scale how well each item measured the survey construct located at the top of the page. The same set of survey items were evaluated against all three construct definitions.

The following instructions were given at the top of each page:

> Rate on a scale from 1-5 how well the survey items measure the survey construct above. A "1" indicates that that the survey item does not measure the construct at all while a "5" indicates that the survey item completely measures the construct.

The final page of the validity survey contained demographic questions about the participants. This included information such as age, gender, and year in school. This information was included to enable post hoc testing to check for participant bias.

The final page of the validity survey also contained two questions which allowed the researcher to determine if the survey participants might have had difficulty completing the validity survey. The first question asked if the participant was a native English speaker. There was a concern that a person who is learning English as a second language may be less able to evaluate the degree to which the items fit within the conceptual construct definitions. The second question asked if the participant had any experience or had taken a class in lean manufacturing. Since survey items contained some terminology associated with lean manufacturing, this question would allow for post-hoc testing to assess whether or not the validity survey results might be impacted by previous exposure to lean manufacturing.

3.2 IRB Approval

Since the survey was completed by human participants, it was necessary to get institutional approval from the Oregon State University Institutional Review Board (IRB). The purpose of obtaining IRB approval is to ensure that human participants are not harmed as a result of the research. The stamp of IRB approval is located on the lower left corner of the survey. (See Appendix A.)

3.3 Data Collection Details

The survey was administered to 35 university students in a lean manufacturing class at Virginia Tech in April of 2006 and 88 university students in a business course at Oregon State University in May of 2006. It was later learned that the definition for the construct, goal difficulty, was incorrect on the Virginia Tech surveys. For this reason, the data from these surveys was not used. The error was fixed before the survey was administered to students at Oregon State University.

The surveys were administered at the end of class to enable students not wishing to participate to leave. A debriefing statement was read to the students by the survey administrator before the survey was distributed. The statement provided information on the purpose of the survey as well as detailed instructions. A copy of this statement is included in Appendix B.

3.4 Participants

University students were used as the target participants consistent with recommendations made by previous researchers such as Hinkin and Tracey (1999). In addition, university students were also the most accessible group for the researcher. Hinkin and Tracey (1999) explain that college students are an excellent choice for content validation studies since they have no “pertinent biases” and have “sufficient intellectual ability to rate the correspondence between items and definitions of various theoretical constructs” (p. 179).

3.5 Data Collection and Analysis

After the data from the surveys was collected, it was entered manually into Microsoft Excel XP by the researcher. The results from each survey page were entered into a separate Excel spreadsheet. In order to ensure that the data could all be traced back to the original paper source if necessary, the researcher assigned each survey a number. The number was written in the upper left hand corner of the survey. The surveys given at Virginia Tech were numbered from 1-35 and the surveys given at Oregon State University were numbered from 36-123.

Data was entered in with the number of the participant in the first column and the code for the item in the row at the top of the Excel spread sheet. The code for each item indicated the construct that the item was measuring. For example, the first item that was used to measure the construct Goal Clarity was labeled GC1. These codes are the same ones that were used by the NSF research team. Table 3.5.1 indicates the code of each item, the item number from the validity survey, and the item statement.

Table 3.5.1: Codes of Item Statements.

Code	**Item #**	**Item Statement:**
GC2	1.	The performance targets our team must achieve to fulfill our goals are clear.
ACC1	2.	In general, members of our team believe in the value of this Kaizen event.
GDF1	3.	Our team’s improvement goals are difficult.
GDF7	4.	Our team has enough time to achieve our goals.
ACC5	5.	Most of our team members think that things would be better without this Kaizen event.
GC4	6.	Our entire team understands our goals.
GDF2	7.	Meeting our team’s improvement goals will be tough.
ACC6	8.	In general, members of our team believe that this Kaizen event is needed.
GDF6	9.	It will take a lot of thought to achieve our team's goals.
GDF4	10.	It will be hard to improve this work area enough to achieve our team’s goals.
ACC4	11.	Most of our team members think that this Kaizen event will serve an important purpose.
ACC2	12.	Most of our team members think that this Kaizen event is a good strategy for this work area.
ACC3	13.	In general, members of our team think that it is a mistake to hold this Kaizen event.
GC1	14.	Our team has clearly defined goals.
GDF3	15.	It will take a lot of skill to achieve our team’s improvement goals.
GC3	16.	Our goals clearly define what is expected of our team.
GDF5	17.	It will take a lot of effort to achieve our team's goals

While entering the data, the researcher double checked the results of one participant after entering data for every five participants. The researcher also went back and checked the data for random participants once all the data had been entered. There were very few errors found, and therefore the results of any errors while entering the data is considered negligible.

Once the data was entered into Excel, the results for pages 2-4 of the validity survey which compared the construct definitions to the survey items were copied and pasted into three separate SPSS 14.0.2 files. Factor analysis was conducted on each data set by specifying the extraction and rotation method for the data. Data that was missing was excluded from the analysis. The number of factors extracted was determined by the number of eigenvalues that were greater than one. The same analyses were also conducted by setting the number of factors extracted to two. Results from both factor analyses were used.

The Wilcoxon signed ranks test was also completed using the software, SPSS 14.0.2. The data from pages 2-4 of the survey was copied and pasted from the Excel spreadsheets into one data file. Each of the survey items was labeled using the construct definition at the top of the page and the code. For example, the results from the first goal clarity item that had been compared to the goal clarity definition were labeled GC_GC1. The results from the first goal clarity item which were compared to the definition of the construct, goal difficulty, were labeled GDF_GC1.

4. RESEARCH FINDINGS

The demographic data was analyzed using basic calculations in Microsoft Excel XP. The data relating the survey constructs to the survey items was analyzed using factor analysis and the Wilcoxon signed ranks test. Both tests showed that the items from the survey were validly measuring the survey construct that they were designed to measure.

4.1 Demographic Research Findings

The number of respondents for which the data was analyzed was 88. The average age of the respondent was 22.6. The majority of the respondents were male (~63%) and almost 90% of the respondents were in their Junior or Senior year of college. The majority of the students (~90%) spoke English as a native language and most (~81%) did not have experience with lean manufacturing.

4.2 Factor Analysis

The first method used to analyze the data was factor analysis. Factor analysis is one of the most common methods used to evaluate the relationships between survey constructs and survey items (Hinkin, 1998). It is an analytical method that simplifies data by evaluating the correlations between variables (Kline, 1994). These correlations are represented by factors (Kline, 1994). In a survey, each factor is essentially a survey construct, and each variable is the item that is being used to measure the factor.

The strength of the relationship between the factor (survey construct) and the variable (survey item) is measured by a factor loading. Factor loadings are represented by a number ranging between 0 and 1. Factor loadings that are less than 0.3 are considered low and can generally be ignored (Kline, 1994). A factor loading that is greater than 0.6 is considered to be high and factors between 0.3 and 0.6 are considered to be moderately high (Kline, 1994). Therefore, items measuring a certain construct should have a high factor loading on one factor, and should have relatively low factor loadings on other factors.

The type of factor analysis used was the principal components method. The matrix was rotated using a Varimax rotation with Kaiser normalization. This is the same method used by Hinkin and Tracey (1999) in their study of content validity. Hinkin (1998) recommends using an orthogonal rotation rather than an oblique rotation because the goal is to have survey items that are "independent of one another" (p. 112).

Table 4.2.1 shows the factor analysis of the items when comparing each of the items to the definition for goal clarity. The number of components to extract was determined by evaluating the number of eigenvalues that were greater than one. Factor loadings lower than 0.3 were suppressed because Kline indicates that these loadings are unimportant (Kline, 1994).

Table 4.2.1: Rotated component matrix for item ratings when compared to the definition for goal clarity.

	Factor			
	1	2	3	4
GDF5	.878			
GDF2	.867			
GDF6	.775			
GDF3	.759			
GDF4	.673			.515
GDF1	.616			
GDF7	.532			.419
ACC2		.871		
ACC6		.866		
ACC4		.864		
ACC1		.707		
GC1			.857	
GC3			.842	
GC4			.745	
GC2			.718	
ACC5				.854
ACC3				.734

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

The table shows that all the items for goal clarity (GC1-GC4) load highly on only one factor. This indicates that the items for goal clarity are correlating very highly with only one factor and not with any of the other factors. This is a good indication that the items used for goal clarity have a high degree of content validity.

One also notices that the items for goal difficulty and most of the items for affective commitment to change load significantly on separate factors. Hinkin indicates that items measuring different survey constructs should load on different factors (Hinkin, 1998). This makes sense because these items are used to measure the same construct definition and thus should correlate higher with one another than with the items for goal clarity.

However, two of the items for affective commitment to change, ACC3 and ACC5, loaded on a separate factor. After reading these two items, it becomes apparent why they would correlate separately from the other ACC items. Item ACC3 states, "In general, members of our team think that it is a mistake to hold this Kaizen event." Item ACC5 states, "Most of our team members think that things would be better without this Kaizen event." Both of these statements are not about the need for a commitment for change but rather about the lack of need for a commitment to change; therefore, the reader probably rated these items more similarly than they would rate the other items for affective commitment to change.

It is also important to keep in mind that although the items for each factor should load separately, Table 4.2.1 was only a comparison of the items to the definition for goal clarity. Therefore, it is most important in this factor analysis that the items for goal clarity load separately than the other items. A plot of the rotated component matrix that is found when only two factors are extracted also makes it clear that the goal clarity items are separate from the other items (see Fig. 4.2.1).

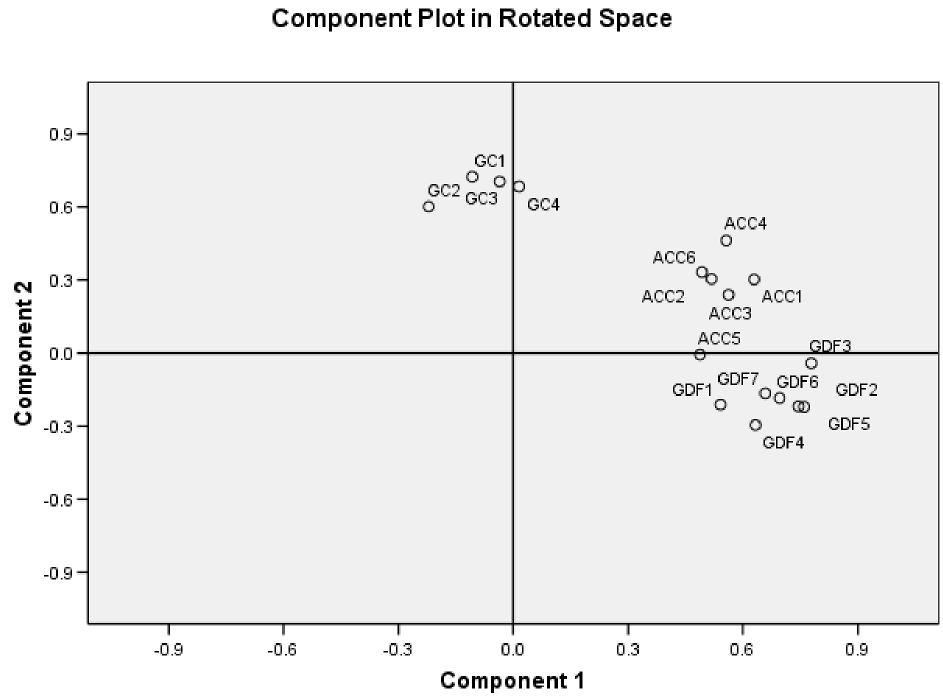

Figure 4.2.1: Plot of rotated components in space of only two components shows that goal clarity items load separately than all the other items.

Factor analysis was also conducted on the results of the items which were compared to the survey construct, goal difficulty. This time, there were three components which had eigenvalues greater than one and three factors were extracted. The results are shown in Table 4.2.2. Six of the items for goal difficulty load highly (> 0.6) on one factor and one loads moderately high (> 0.3). The item loading moderately high, however, is very close to 0.6 at 0.564.

Table 4.2.2: Rotated component matrix for item ratings when compared to the definition for goal difficulty.

	Factor		
	1	2	3
GC1	.885		
GC3	.843		
GC4	.811		
GC2	.788		
ACC2	.622		.448
ACC4	.568	-.311	.560
GDF6		.788	
GDF2		.785	
GDF5		.732	
GDF1	-.424	.728	
GDF4	-.402	.708	
GDF3		.637	-.360
GDF7		.564	
ACC3			.824
ACC5			.801
ACC6	.449		.713
ACC1	.557		.573

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

Two of the items for goal difficulty also cross load moderately high on the first factor and one of the items cross loads moderately highly on the third factor. Reading the items does not provide any insight into why the items load this way. However, the cross loadings are significantly less than the loading for the second factor, and a plot of the rotated component matrix for two items also show that the goal difficulty items are separate from the other items (See Fig. 4.2.2).

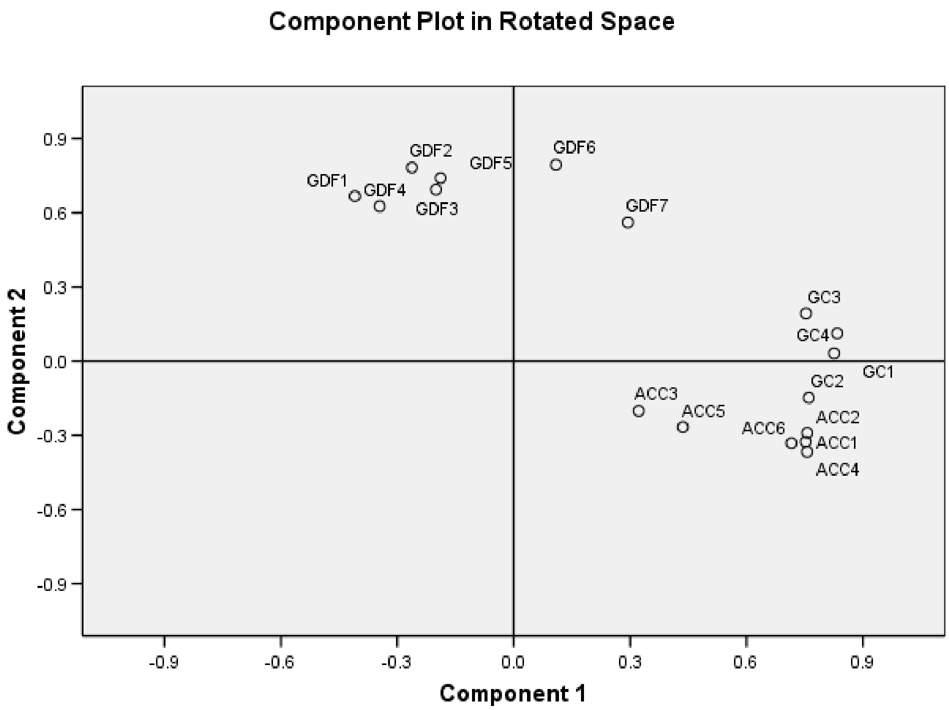

Figure 4.2.2: Plot of rotated components in space of only two components shows that items for goal difficulty load separately than all the other items.

The other items in Table 4.2.2 load on separate factors than the items for goal difficulty. All the goal clarity items load on one factor, but the items for affective commitment to change once again load on separate factors. All of these items load on factor three, but four of the items load moderately high on factor one. This is not particularly significant since these items were being compared to the definition for goal difficulty.

Table 4.2.3 shows the results for the factor analysis of the items compared to the definition for affective commitment to change. Four of the items load highly (>0.6) on one factor. The other two items load moderately high on this factor but they also cross load very highly on the second factor.

Table 4.2.3: Rotated component matrix for item ratings when compared to the definition for affective commitment to change.

	Factor		
	1	2	3
GDF3	.829		
GDF2	.808		
GDF1	.804		
GDF5	.802		
GDF4	.787		
GDF6	.764		
GC1	.315	.788	
GC3	.312	.756	
ACC3		-.744	.398
GC4	.396	.730	
ACC5		-.716	.434
GC2		.692	
GDF7	.557	.582	
ACC6			.881
ACC4			.863
ACC2			.862
ACC1			.826

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

The two items that cross load on factor two are ACC3 and ACC5. As mentioned earlier these two items are different than the other items for affective commitment to change. They are negative statements about the lack of need for change. The definition used for the ACC construct indicates that corresponding items should "describe team perceptions of the need for the specific changes targeted by the Kaizen event." The two items which did not correlate highly with this definition are negative statements about how unnecessary the Kaizen events would be. They do not talk about the "need" for the changes but rather are about the lack of "need". Since this is the case, it is understandable why these two items cross-correlate with the other factors.

A plot of the rotated component matrix using two factors shows that the items for the affective commitment to change construct are much more highly correlated with one another than with the other items (see Fig. 4.2.3). The two negative statements also are shown to be closer to one another and a little bit further from the other items for the affective commitment to change construct.

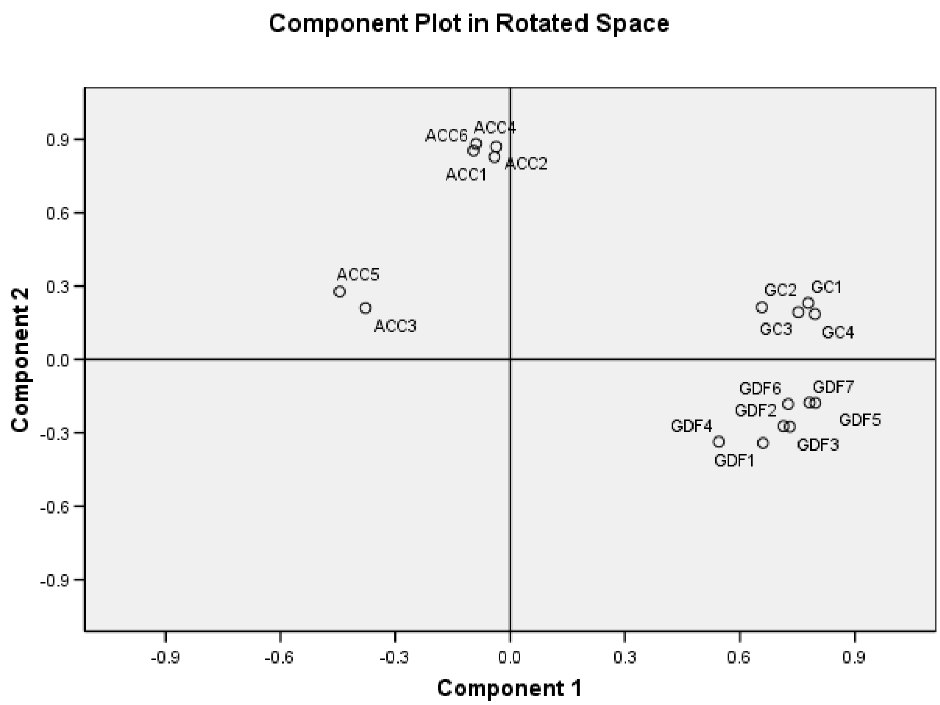

Figure 4.2.3: Plot of rotated components in space of only two components shows that items for the affective commitment to change construct load differently than all the other items.

The results of the factor analysis show that the majority of items are loading either moderately high or highly on one factor and either not very high or not at all on other factors. This is evidence of content adequacy which is similar to content validity (Hinkin & Tracey, 1999). This is also an indicator of construct validity (Hinkin & Tracey, 1999).

The factor analysis method used is also similar to exploratory factor analysis. Usually, exploratory factor analysis is conducted while creating the survey to help determine whether or not the items are validly measuring the construct they were designed to measure (Hinkin, 1998). This is usually completed as the second step for establishing construct validity (Serafini & Adams, 2002; Hinkin, 1998). In this investigation, the factor analysis was completed after the kaizen event kick-off survey was already finalized, however, the method was still the same and the factor loadings still indicate that the survey items loaded highly on the constructs they were designed to measure. These results can be used as evidence of construct validity.

4.3 Wilcoxon Signed Ranks Test

Although factor analysis is a common method used to evaluate the relationships between survey items and survey constructs, it is a somewhat subjective analytical technique. Hinkin and Tracey (1999) consider it to be a risky method to use when establishing validity because many subjective decisions need to be made to complete it.

Hinkin and Tracey (1999) recommend using another method to help determine if an item should or should not be used to evaluate a particular survey construct. They recommend comparing the mean rating of the item for its proposed construct definition to the mean rating of the item on the other two definitions. If the mean rating for the item on the construct it was designed to evaluate is significantly higher than it is when compared to the other construct definitions, then that item is "consistent with the proposed theoretical construct" (Hinkin and Tracey, 1999)

To compare the means between items, Hinkin and Tracey (1999) used a one-way analysis of variance (ANOVA). They argued that ANOVA could be used because it is lenient towards distributions that are not entirely normal.

In the current investigation, the Wilcoxon signed ranks test was used. It was deemed to be more appropriate because it is a non-parametric test, and does not require that the data come from a normal distribution or that the variances be assumed equal (Grimm, 1993). Table 4.3.1 shows the results of the Wilcoxon signed ranks test.

Table 4.3.1: Results of the Wilcoxon signed ranks test.

Item	Test Indicates:	Z Value	P value
GC 1	GC_GC1 > GDF_GC1	-5.305	0.000
	GC_GC1 > ACC_GC1	-6.234	0.000
GC 2	GC_GC2 > GDF_GC2	-4.479	0.000
	GC_GC2 > ACC_GC2	-5.169	0.000
GC 3	GC_GC3 > GDF_GC3	-5.941	0.000
	GC_GC3 > ACC_GC3	-5.877	0.000
GC 4	GC_GC4 > GDF_GC4	-4.797	0.000
	GC_GC4 > ACC_GC4	-5.306	0.000
GDF1	GDF_GDF1 > GC_GDF1	-6.563	0.000
	GDF_GDF1 > ACC_GDF1	-6095	0.000
GDF2	GDF_GDF2 > GC_GDF2	-6.026	0.000
	GDF_GDF2 > ACC_GDF2	-5.668	0.000
GDF3	GDF_GDF3 > GC_GDF3	-3.151	0.000
	GDF_GDF3 > ACC_GDF3	-4.658	0.000
GDF4	GDF_GDF4 > GC_GDF4	-4.244	0.000
	GDF_GDF4 > ACC_GDF4	-4.489	0.000
GDF5	GDF_GDF5 > GC_GDF5	-4.543	0.000
	GDF_GDF5 > ACC_GDF5	-4.784	0.000
GDF6	GDF_GDF6 > GC_GDF6	-4.428	0.000
	GDF_GDF6 > ACC_GDF6	-4.772	0.000
GDF7	GDF_GDF7 > GC_GDF7	-3.840	0.000
	GDF_GDF7 > ACC_GDF7	-4.401	0.000
ACC1	ACC_ACC1 > GC_ACC1	-4.096	0.000
	ACC_ACC1 > GDF_ACC1	-5.379	0.000
ACC2	ACC_ACC2 > GC_ACC2	-5.946	0.000
	ACC_ACC2 > GDF_ACC2	-4.552	0.000
ACC3	ACC_ACC3 > GC_ACC3	-5.269	0.000
	ACC_ACC3 > GDF_ACC3	-4.781	0.000
ACC4	ACC_ACC4 > GC_ACC4	-5.322	0.000
	ACC_ACC4 > GDF_ACC4	-4.202	0.000
ACC5	ACC_ACC5 > GC_ACC5	-5.051	0.000
	ACC_ACC5 > GDF_ACC5	-4.722	0.000
ACC6	ACC_ACC6 > GC_ACC6	-5.415	0.000
	ACC_ACC6 > GDF_ACC6	-4.524	0.000

Table 4.3.1 shows that all the items were evaluated with significantly higher ratings when compared to their own construct definition than they did when compared to another construct definition. This test is not as subjective as the factor analysis and shows that all items in the survey demonstrate content validity (Hinkin and Tracey, 1999).

4.4 Recommendations

While the second test shows that all the items have content validity, the first test does not indicate that all the items necessarily have construct validity. The items ACC3 and ACC5 both did not show construct validity primarily because the items were negatively worded when compared to other items that measure the affective commitment to change construct. In order to evaluate if these items have construct validity, it would be necessary to revise the definition for affective commitment to change to evaluate both positively and negatively worded items; however, since the affective commitment to change construct already has six items, it would be easier to simply delete both items from the study. Only four items are needed to evaluate this construct and thus deleting the items would probably improve the results (Hinkin, 1998).

The other item which did not appear to have a high degree of construct validity was GDF7. It is recommended that this item also be deleted from the study because the construct, goal difficulty, already has seven items which is well above the amount needed to validly measure a construct (Hinkin, 1998).

4.5 Conclusion

All of the items except for ACC3, ACC5, and GDF7 loaded highly on one factor when compared with the construct that they were designed to measure. This indicates that these items demonstrate construct validity. The fact that the majority of the items loaded primarily on one factor also indicates that the items have high content validity.

The comparison of means using the Wilcoxon signed ranks test also gave evidence for content validity. All of the tests were significant which indicates that the signed ranks statistic for each of the items was much greater when compared to the construct definition that they were designed to measure than when compared to the other construct definitions. This is an excellent indication of content validity which is also a necessary part of construct validity (Hinkin & Tracey, 1999).

Bibliography:

Babbie, E. (1998). The Practice of Social Research. (8th ed.). Westford, MA: Wadsworth Publishing Company.

Black, J.T. & Hunter, S.L. (2003). Lean Manufacturing Systems and Cell Design. Dearborn, MI: Society of Manufacturing Engineers.

Chow, W.S. & Lui, K.H. (2003). A structural analysis of the significance of a set of the original TQM measurement items in information systems function. Journal of Computer Information Systems, 43(3), 81-91.

Day, J.C. (1998). Learning About Lean Systems at Freudenberg-NOK: Where Continuous Improvement Is a Way of Life. In J.K. Liker (Ed.), Becoming Lean (pp. 179-198). Portland, OR: Productivity Press.

Grimm, L.G. (1993) Statistical Applications for the Behavioral Sciences. United States: John Wiley & Sons, Inc.

Hinkin, T.R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. Organizational Research Methods, 1(1), 104-121.

Hinkin, T.R. & Tracey, J.B. (1999). An analysis of variance approach to content validation. Organizational Research Methods, 2(2), 175-186.

Kline, P. (1994). An Easy Guide to Factor Analysis. New York, NY: Routledge.

Liker, J.K. (1998). Conclusion: What We Have Learned about Becoming Lean. In J.K. Liker (Ed.), Becoming Lean (pp. 497-517). Portland, OR: Productivity Press.

McGrath, R.E. (2005). Conceptual complexity and construct validity. Journal of Personality Assessment, 85(2), 112-124.

Muchinsky, P.M. (1997). Psychology Applied to Work. (5th ed.). Pacific Grove, CA: Brooks/Cole Publishing Company.

Nicholas, J.M. (1998). Competitive Manufacturing Management. United States of America: The McGraw-Hill Companies, Inc.

Ortiz, C. (2006). All-out kaizen. Industrial Engineer, 38(4), 30-34.

Rother, M. (1998). Crossroads: Which Way Will You Turn on the Road to Lean. In J.K. Liker (Ed.), Becoming Lean (pp. 477-495). Portland, OR: Productivity Press.

Serafini, T.E. & Adams, G.R. (2002). Functions of identity: scale construction and validation. Identity: An International Journal of Theory and Research, 2(4), 361-389.

Shook, J.Y. (1998). Bringing the Toyota Production System to the United States: A Personal Perspective. In J.K. Liker (Ed.), Becoming Lean (pp. 41-69). Portland, OR: Productivity Press.

Womack, J.P. & Jones, D.T. (1996). Lean Thinking. New York, NY: Simon & Schuster.

Yusof, S.M. & Aspinwall, E.M. (2000). Critical success factors in small and medium enterprises: survey results. Total Quality Management, 11, 448-462.

APPENDICES

APPENDIX A

VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY

Dear Student,

Your help is needed for important research! Dr. Toni Doolen, a professor in the Department of Industrial and Manufacturing Engineering, is working on a research project sponsored by the National Science Foundation (NSF) to determine how to design a kaizen event in order to make it successful. A kaizen event is a short-term project focused on improving a specific process or set of activities, such as the work flow within a specific work center. One of the tools developed to evaluate the effectiveness of kaizen events is a kickoff survey. This survey is given to employees at the beginning of a kaizen event and is used to evaluate how different employee attitudes might affect event outcomes.

The research team is interested in testing the validity of the kickoff survey. Language can be subjective, and the research team must determine if the surveys are actually measuring what they were designed to measure. Rebecca Mitchell, an undergraduate student in the Department of Industrial and Manufacturing Engineering, is leading this research project to study the validity of the kaizen event kickoff survey. Your participation is being requested to help evaluate whether or not the survey items developed for the research match the concepts that the survey is measuring (these concepts are called survey constructs).

The survey should take about 15-20 minutes of your time. Completing the survey is voluntary. There are no foreseeable risks or direct benefits to participants who participate in this survey. If you are taking the survey, there will be no penalty to your grade or your standing in the university if you choose not to participate. Do not participate in the survey if you are under 18. Do not include your name or any other identifier on the survey. The confidentiality of your responses will be protected to the extent permitted by law.

If you wish to participate, please respond to the questions starting on the next page. For each page of the study, you will be asked to review a survey construct and its corresponding definition. After reading the construct and definition, review each of the 17 items listed below and rate how well each item measures the concept defined by the survey construct. A rating of "1" indicates that that the survey item does **not** relate to the construct definition at all while a rating of "5" indicates that the survey item relates completely to the construct definition. You may decline to rate any of the items.

Thank you for your help in this important research! If you have any questions or comments, please contact Rebecca Mitchell, mitchere@onid.orst.edu. If you have any questions about your rights as a research subject, please contact the Oregon State University Institutional Review Board (IRB) Human Protections Administrator, OSU Research Office, (541)737-4933, IRB@oregonstate.edu.

Please detach this cover letter from the survey and keep it for your records.

OSU IRB Approval Date: 3-13-06
Approval Expiration Date: 3-12-07

Survey Construct
Goal Clarity: These items describe team perceptions of the extent to which the kaizen event team's improvement goals have been clearly defined.

Instructions
Rate on a scale from 1-5 how well the survey items measure the survey construct above. A "1" indicates that that the survey item does not measure the construct at all while a "5" indicates that the survey item completely measures the construct.

Items:	Not at all	A little bit	Somewhat	Mostly	Completely
1. The performance targets our team must achieve to fulfill our goals are clear.	1	2	3	4	5
2. In general, members of our team believe in the value of this Kaizen event.	1	2	3	4	5
3. Our team's improvement goals are difficult.	1	2	3	4	5
4. Our team has enough time to achieve our goals.	1	2	3	4	5
5. Most of our team members think that things would be better without this Kaizen event.	1	2	3	4	5
6. Our entire team understands our goals.	1	2	3	4	5
7. Meeting our team's improvement goals will be tough.	1	2	3	4	5
8. In general, members of our team believe that this Kaizen event is needed.	1	2	3	4	5
9. It will take a lot of thought to achieve our team's goals.	1	2	3	4	5
10. It will be hard to improve this work area enough to achieve our team's goals.	1	2	3	4	5
11. Most of our team members think that this Kaizen event will serve an important purpose.	1	2	3	4	5
12. Most of our team members think that this Kaizen event is a good strategy for this work area.	1	2	3	4	5
13. In general, members of our team think that it is a mistake to hold this Kaizen event.	1	2	3	4	5
14. Our team has clearly defined goals.	1	2	3	4	5
15. It will take a lot of skill to achieve our team's improvement goals.	1	2	3	4	5
16. Our goals clearly define what is expected of our team.	1	2	3	4	5
17. It will take a lot of effort to achieve our team's goals	1	2	3	4	5

OSU IRB Approval Date: 3-13-06
Approval Expiration Date: 3-12-07

Survey Construct
Goal Difficulty: These items describe team perceptions of the difficulty of the improvement goals set for the Kaizen event team.

Instructions
Rate on a scale from 1-5 how well the survey items measure the survey construct above. A "1" indicates that that the survey item does not measure the construct at all while a "5" indicates that the survey item completely measures the construct.

Items:	Not at all	A little bit	Somewhat	Mostly	Completely
1. The performance targets our team must achieve to fulfill our goals are clear.	1	2	3	4	5
2. In general, members of our team believe in the value of this Kaizen event.	1	2	3	4	5
3. Our team's improvement goals are difficult.	1	2	3	4	5
4. Our team has enough time to achieve our goals.	1	2	3	4	5
5. Most of our team members think that things would be better without this Kaizen event.	1	2	3	4	5
6. Our entire team understands our goals.	1	2	3	4	5
7. Meeting our team's improvement goals will be tough.	1	2	3	4	5
8. In general, members of our team believe that this Kaizen event is needed.	1	2	3	4	5
9. It will take a lot of thought to achieve our team's goals.	1	2	3	4	5
10. It will be hard to improve this work area enough to achieve our team's goals.	1	2	3	4	5
11. Most of our team members think that this Kaizen event will serve an important purpose.	1	2	3	4	5
12. Most of our team members think that this Kaizen event is a good strategy for this work area.	1	2	3	4	5
13. In general, members of our team think that it is a mistake to hold this Kaizen event.	1	2	3	4	5
14. Our team has clearly defined goals.	1	2	3	4	5
15. It will take a lot of skill to achieve our team's improvement goals.	1	2	3	4	5
16. Our goals clearly define what is expected of our team.	1	2	3	4	5
17. It will take a lot of effort to achieve our team's goals	1	2	3	4	5

OSU IRB Approval Date: 3-13-06
Approval Expiration Date: 3-12-07

Survey Construct
Affective Commitment to Change (Commitment to Event Goals): These items describe team perceptions of the need for the specific changes targeted by the kaizen event.

Instructions
Rate on a scale from 1-5 how well the survey items measure the survey construct above. A "1" indicates that that the survey item does not measure the construct at all while a "5" indicates that the survey item completely measures the construct.

Items:	Not at all	A little bit	Somewhat	Mostly	Completely
1. The performance targets our team must achieve to fulfill our goals are clear.	1	2	3	4	5
2. In general, members of our team believe in the value of this Kaizen event.	1	2	3	4	5
3. Our team's improvement goals are difficult.	1	2	3	4	5
4. Our team has enough time to achieve our goals.	1	2	3	4	5
5. Most of our team members think that things would be better without this Kaizen event.	1	2	3	4	5
6. Our entire team understands our goals.	1	2	3	4	5
7. Meeting our team's improvement goals will be tough.	1	2	3	4	5
8. In general, members of our team believe that this Kaizen event is needed.	1	2	3	4	5
9. It will take a lot of thought to achieve our team's goals.	1	2	3	4	5
10. It will be hard to improve this work area enough to achieve our team's goals.	1	2	3	4	5
11. Most of our team members think that this Kaizen event will serve an important purpose.	1	2	3	4	5
12. Most of our team members think that this Kaizen event is a good strategy for this work area.	1	2	3	4	5
13. In general, members of our team think that it is a mistake to hold this Kaizen event.	1	2	3	4	5
14. Our team has clearly defined goals.	1	2	3	4	5
15. It will take a lot of skill to achieve our team's improvement goals.	1	2	3	4	5
16. Our goals clearly define what is expected of our team.	1	2	3	4	5
17. It will take a lot of effort to achieve our team's goals	1	2	3	4	5

OSU IRB Approval Date: 3-13-06
Approval Expiration Date: 3-12-07

The following questions will not be used for identification purposes. All responses are voluntary.

Please fill in the blank.

1 **Age: ________**

Please circle the appropriate response for each of the following questions.

2 **Sex:**

Male Female

3 **Year in School:**

Freshman-Sophomore Junior-Senior Graduate Student

4 **Is English your first language?**

Yes No

5 **Have you had any experience with lean manufacturing or have you taken a class on lean manufacturing at OSU or elsewhere?**

Yes No

OSU IRB Approval Date: 3-13-06
Approval Expiration Date: 3-12-07

APPENDIX B

Recruitment Material

This document will be used to recruit participants. The text will be read to students in classes (with the permission of the instructor) to get volunteers to participate in the study. If additional participants are needed, this text will be e-mailed to potential participants (undergraduate and graduate students in the Colleges of Engineering and Business at OSU).

Dear Student,

Your help is needed for an important research study! Dr. Toni Doolen, a professor in the Department of Industrial and Manufacturing Engineering, is working on a research project sponsored by the National Science Foundation (NSF) to determine how to design successful kaizen events in manufacturing companies. A kaizen event is a short-term project focused on improving a specific process or set of activities, such as the work flow within a specific work center. One of the tools used by the research team to evaluate the effectiveness of kaizen events is a kickoff survey. This survey will be given to the employees at the beginning of a kaizen event and will be used to evaluate how different employee attitudes affect the outcomes of a kaizen event.

The research team is interested in testing the validity of the kickoff survey. Language can be subjective, and the research team must determine if the surveys are actually measuring what they were designed to measure. Rebecca Mitchell, an undergraduate student in the Department of Industrial and Manufacturing Engineering, is leading this research project to study the validity of the kaizen event kickoff survey. Your participation is being requested to help evaluate whether or not the survey items developed for the research match the concepts that the survey is measuring (these concepts are called survey constructs).

We are looking for volunteers to complete the validation of the survey. The study will take approximately 15-20 minutes to complete. No previous experience with kaizen events is needed to participate in this study. If you choose to participate you are not required to provide your name, so your participation will be anonymous. You may not participate if you are under 18. If you have any questions or comments, you may contact Rebecca Mitchell at mitchere@onid.orst.edu. If you have any questions about your rights as a research subject, please contact the Oregon State University Institutional Review Board (IRB) Human Protections Administrator, OSU Research Office, (541)737-4933, IRB@oregonstate.edu.

Thank you.

Rebecca Mitchell, Industrial and Manufacturing Engineering
Dr. Toni L. Doolen, Assistant Professor, Industrial and Manufacturing Engineering