

Improving Safety in a High Reliability/Low Commitment Work Environment

K.J. McNamara, J.M. Thom, and R.E. Thompson
Purdue University

ABSTRACT

Over the years several strategies were used to create safe and productive work environments. While all of these methods made an impact on the employees and the work environments, most results were short lived and in some instances created a reverse effect that actually made the employees less productive and less safe. To be successful, it was important for the employees to “buy-in” to these policies and procedures, and to “understand, accept and appreciate” them (Geller, 2001). When this occurred, behaviors would begin to change and in turn, the culture of the work group or organization began to change as well. The measurement of employees’ perceptions regarding the acceptance level of a possible training classification program in previous studies at Purdue University was positive. Based on this historical data a new study was done at a major U.S. air carrier in an effort to gather information regarding employees’ views and opinions on the possible implementation of a system of color coding to identify the job classification of the employees working in an environment requiring high reliability, and with a low commitment to the job, such as part time employees. Based on the findings of this study, it was concluded that a training classification system based on color coding could be accepted and supported by low commitment, high reliability organizations. The name of the company studied was withheld in this paper due to proprietary considerations, and was identified in this study as *Company X*.

INTRODUCTION

Companies struggle with the necessity of providing a safe and productive work environment for their employees. The complexity of daily work operations and the overwhelming number of safety issues may leave workers without the proper tools and resources to accomplish their jobs. Research into this problem has led to new programs to change the work environment’s culture by changing individual’s behavior patterns (McSween, 1995). An area of concentration in this research has focused specifically in workplace safety. Preliminary results are positive when the programs are implemented within stable, full-time work groups. It is the purpose of this study to apply the theories behind this new research and test whether these programs can also work for similar high reliability work environments utilizing as their primary work force part-time and temporary workers with high turnover rates. This study focuses on the implementation of a training classification system within such a work group at a major U.S. air carrier, which is referred to in this study as *Company X*.

BACKGROUND

Over the years strategies have been used to create safe and productive work environments. These included, but were not limited to: social pressure (Asch, 1995), obedience (Milgram, 1963), negative reinforcement (Endler & Hartley, 1973), and positive reinforcement (Koepnick, 1993). While all of these methods made an impact on the employees and the work environments, most results have been short lived and in some instances created a reverse effect that actually made the employees less productive and less safe.

Within the last decade, three prominent and respected researchers, James Reason, Scott Geller, and Terry McSween, have narrowed their research focus to perfecting behavioral approaches that help to create safe and productive workplace environments. Similarity is found in their work, based on focus of the interaction of people with themselves, other people, and with the work environment. Reason has focused on the “defenses” or safeguards organizations can put in place so that accidents may be prevented (Reason, 2000). Geller has concentrated on behavioral approaches to safety and productivity, and attempts to share the

reasoning behind why these types of programs are so effective (Geller, 2001). Finally, McSween goes one step further by detailing how to integrate behavioral systems into the workplace (McSween, 1995). All of these authors have categorized training classification systems as an approach aimed at changing the behavior and/or culture of the work group, by changing the behavior of the individual employees and their interactions with each others and their work environment.

Reviewing Reason, Geller, and McSween, the theory behind why behavioral based programs are so effective becomes clear. Behavior based programs “develop a set of comprehensive principles on which to base safety procedures and policies” (Geller, 2001, 21). It is important for the employees to “buy-in” to these policies and procedures and “understand, accept, and appreciate” them (Geller, 2001). When this occurs, behaviors begin to change and in turn, the culture of the work group or organization begins to change as well. The method of the current study closely follows the ideas and information found in Thomas Krause’s book, *The Behavior-Based Safety Process* (1997). Krause (1997) has proposed to use the following implementation sequence with behavior-based programs (p. 95):

1. Implementation Planning Meeting,
2. Assessment Visit and Report,
3. Behavioral Inventory Tools Development, Management Training, Ownership Meetings,
4. Observation Course Development,
5. Observer Course Review, Computer Software Training, Observer Training, Kickoff Meetings,
6. Ongoing Observations and Data collection, Process Checks,
7. Safety Improvement Process Training,
8. Action Planning, Users Conferences, Benchmarking.

Prior study and research (Hess, 2000) has covered steps one through five. The training classification study done here focuses on steps six and eight.

A review of literature on training classification systems reveals several programs

already in place. The United States Navy has long used a classification system program on aircraft carrier flight decks. Working in such a hazardous environment, verbal communication is nearly impossible due to the high levels of noise and number of tasks being performed simultaneously. In light of these difficulties, a highly evolved set of hand signals and color-coded vests have been put into place. (Paige, 1998). Each person on the flight deck has a specific function. There are fireguards, fuelers, pilots, mechanics, flight deck officers, and deck edge officers, just to name a few. Each job classification is assigned a specific color vest, and everyone must know what each color means prior to being allowed onto the flight deck. Both the operations on the aircraft carrier flight decks and the operations on the aircraft ramp of *Company X* fit the description of a high-reliability organization (HRO). Reason states, “Organizational flexibility means possessing a culture capable of adapting effectively to changing demands. Flexibility is one of the defining properties of...high-reliability organizations (HROs)” (2000, 213). Basically, the amount of critical job responsibilities and the ever changing environment in which they must be performed classify *Company X* workers to be called a high-reliability organization. Within an HRO there are many operational challenges occurring during high demand and peak production periods (Reason, 2000). Utilizing a training classification system was expected to increase employees’ awareness of their surroundings by providing additional means (color-coded clothing) for assessment.

Initial work between Purdue University and *Company X* began in the fall of 1998. Some of the data and information collected during this time was used throughout the course of this current study. The measurement of the employees’ perceptions regarding the acceptance level of a possible training classification system program was positive (Hess, 2000). Based on this response, focus groups were held in an effort to gather information regarding employees’ views and opinions on the possible implementation of a program of this kind. The focus group questions solicited information regarding; the benefits of a training classification program, the potential

problems implementation may cause, and the anticipated employee support and acceptance levels. Once again results proved that a training classification program might have been a viable option *Company X* should explore (Hess, 2000).

There have been examples of other forms of employee classification programs that have been accepted by employees, and have proved to increase employee situational and safety awareness. For example, Southwest Airlines CEO, Herb Kelleher decided to “code” employees by their personality types. By identifying certain personality traits, employees had a better understanding of how to approach each other and how to interpret individual responses (Freiberg & Freiberg, 1996). Within a month after implementation, there was shown to be a drop in safety violations as observed by an independent consultant (Freiberg & Freiberg, 1996). In addition, Southwest Airlines prided itself on fostering a family and team culture. The employee classification system reinforced to the employees the airline’s commitment to this value.

Personal injuries and equipment damage cost *Company X* millions of dollars each year¹. Surprisingly, the dollar amounts only include the treatment of injuries and the cost of parts to make repairs to the equipment. The data has not included lost time from work, overtime wages to fix damaged equipment, or lost revenues because the company does not have a system to track these costs. This spending is not budgeted into a separate account, because the severity and number of accidents is not seen as predictable from year to year. However, now that *Company X* reports a drop in quarterly profits, efforts have been made to cut spending in several areas (Connor, 2001). The information regarding personal injuries and equipment damage obviously has the attention of high-level officers in the company and is seen as a target area to reduce spending. Efforts have been taken to find solutions to decrease the amount of personal injuries and equipment damage. This study, being one of those efforts, attempts to show that employees can accept new programs based on

behavior changing strategies at this *Company X* station.

PROCEDURES

Performing research in an industrial setting posed certain challenges and had limitations not present in a laboratory or controlled setting. In this particular case, due to a limited number of observers, the observations were limited to only two of the thirteen aircraft offload teams. In addition, the study was conducted at only one *Company X* station. Although permission was granted to perform the study, *Company X* did reserve the right to monitor and/or control the release of the information generated by the study if deemed necessary.

The intent of this study was to determine in a high turnover environment, where high reliability was needed, whether the employees could successfully accept and utilize a simple visual classification system to provide awareness of the levels of skills and experience necessary for certain job activities.

Pre-test and post-test surveys were developed based on the information gathered during the initial focus groups. These surveys were pilot tested by a group of offload workers representing the various types of employees that would be involved in the observations. The workers who participated in the pilot test were not a part of the remainder of the study. Corrections were made to ensure the survey would accurately collect the desired information. Two offload teams under different managers were then classified by using three different color coded armband ID holders. This classification took place during a pre-work meeting held onsite in a pre-designated area. During this meeting, the pre-test survey was administered and the program explained to the employees. The workers then received the proper color armband for their predetermined skill and experience. Each color specified the types of specialized training the employee had completed. Red was used for newly hired employees, blue was given to employees qualified to operate equipment on the ramp, and

¹ *Company X* has deemed actual figures as proprietary and confidential, however they were made available to the researcher.

yellow was assigned to team heads². All employees were required to wear an armband at work in order to show their FAA (Federal Aviation Administration) authorization to be on the flight ramp area. This made implementation very easy to accomplish. The researchers then, using an at-risk behavior checklist already developed for other university research³ (Fought, 2000) observed the two “coded” work groups. After a period of four weeks, all employees in the observed work groups completed a post-test survey to gather information regarding the classification system trial. The post-test survey data was then compared to the pre-test survey data using a Statistica database. An ANOVA (analysis of variance) was performed on all questions in order to determine if there was a significant ($p>0.05$) relationship between the two sets of data (Sekaran, 2000). In addition, a t-test was performed because it was not known if the difference would be positive or negative (Sekaran, 2000). The open-ended questions were entered into the database exactly as written by the employees, and provided to the company for their review (preserving the employees’ anonymity) on the training classification program trial. Evaluation, comparison, and statistical analysis of the observation data using an Access® based program (Lee, 2001) were also completed.

FINDINGS AND DISCUSSION

Upon completion of the training classification system and analysis of the data gathered, several findings were evident. With regard to the survey responses, the ANOVA test indicated there was no significant difference ($p>0.05$) between the answers given on the pre-test versus the answers given on the post-test. More specifically, the ANOVA compared five identical questions from the pre-test and post-test surveys and concluded the answers were the

same. This indicated the employees accepted the training classification system. While the ANOVA test showed a slightly negative correlation among the tested questions, none of the correlations were significant, and therefore it may be concluded the surveys were not significantly different. In addition, two-tailed t-tests were performed on the five repeated questions; however a 5x5 factorial design induced an error rate, making this data unusable. The pre-test survey contained questions aimed at gaining the employees’ perceptions and opinions on the use of a training classification system in their work environment. Most responses were positive and stated that a classification system would be accepted if implemented into policy. The written responses also provided valuable information about the system trial. Many employees reported the armbands made assignment of tasks easier and aided newer employees in the group. Several people also stated that if the armband system was incorporated across the station, it would make operation of a mixed work group (a temporary work group consisting of many employees from different areas) much easier because an employee’s qualifications would be easily visible.

The observation data also supported the acceptance of the classification system.⁴ A total of 36 individual observations were conducted throughout the course of the trial. Referring to Appendix A, a significant drop in the number of safety violations was noted. On the first night of observations, 37 safety violations were observed. Subsequent observations yielded violation numbers of 26, 21, 23, 10, 1, and 9 respectively. These numbers confirmed that a downward trend in respect to safety violations occurred during the observations. Another noted observation was the increase of safety topics brought up in pre-work meetings, as compared to the information gathered from the Safety Perception Survey.

As previously mentioned, several limitations occurred during the course of this study. The

² Colors were chosen based on availability from the armband vendor.

³ Student observers had completed a standard training program for using the safety metrics package at Purdue University. This metric was successfully used at other major air carriers.

⁴ Limitations such as safety audits, manager explanations of the program to employees, observer interactions with the work groups, and company memos may render this data unreliable.

practice of not using new hires on offload teams forced a change in the levels of training classifications in this study. Originally, red was to be used for new hires with no training, blue would be given to employees qualified to operate offload equipment, and yellow would be worn by team leads. In light of the changes, the following classification system was instead used: red was given to employees qualified to drive aircraft tugs, blue was given to fully qualified (tug and belt loader qualified) employees and yellow was worn by the team leads. Although the system was used differently from initially planned, the study was not affected by this change. The completion of an internal safety audit during the system trial period, which resulted in a memorandum instructing managers to “crack down” on safety violations, may have had an affect on the observation data. Extra emphasis given to the importance of “being safe” on the ramp was usually shown as a temporary increase in awareness from the employees. However, during the work operation, when the pressures of “getting the job done” were noticeable, the employees seemed to revert back to normal behaviors. Since the observations of this study focused only on the actual offloading of aircraft, a time when this reversion pressure was present, the effect of short term improvement was noted but believed to be minimal. Another factor that may have affected the study concerns the stability of the work groups. Several employees did not complete both the pre-test and post-test surveys. This fact may have lowered the reliability of the survey data. In order to minimize this effect, employees who entered the system trial after the kick off were briefed on the study, given the appropriate color armband, and had the opportunity to ask questions regarding the study. Finally, the interaction of the observers and the work groups may have affected the observation data. It was evident that some signs of the Hawthorne Effect existed during the study.

The Hawthorne Effect occurs when workers’ job performance improves following the start of the researcher’s intervention (Muchinsky, 1983). Muchinsky stated, “Performance continues to improve because of the novelty of the situation; e.g. the employees

respond positively to the novel treatment they were getting from the researchers. Eventually, however, the novelty begins to wear off, and productivity levels return to their earlier level” (1983, 19). It is important to note, the drop in observed safety violations may be attributed to the attention given to the work groups.

However, perhaps because the observation trial in this study was only four weeks long, there were no signs of safety violations returning to the level noticed at the beginning of the trial. It was found that sometimes behavior changes were due just to a change in the environment. (Muchinsky, 1983). In this case, the observed employees were separated from the rest of the work environment by the different color armband colors and were made aware of the observers’ presence. Throughout the course of the study, several employee participants involved in the study asked questions regarding what the observers were looking for. Giving this information to the employees may have altered the employees’ behaviors.

CONCLUSIONS AND RECOMMENDATIONS

Based on the findings of this study, it is concluded that a training classification system can be accepted and supported by low commitment, high reliability organizations. For this station of *Company X*, it is recommended the training classification system be tested at the entire station level, and further research be conducted in an effort to validate the current study’s findings, and to identify if the armbands could contribute to an increase in safety and/or productivity.

If the assumption is made that the acceptance of the training classification system could be validated station wide, the next logical recommendation would be to test the effects of the classification system on the entire work environment. Identifying the effects of the armbands may be tested by using additional work groups. Steps would have to be taken in order to reduce or eliminate the limitation of the current study. For example, implement the armband system and make no observations during the trial period, or make blind observations of the work groups, so the

employees would have no knowledge they are being observed. Placing observers in the control tower, and scheduling the work groups only to gates where there would be an unobstructed view could accomplish these blind observations. It would, however, require cooperation from the scheduling center. This would reduce the possible problem with the Hawthorne Effect. It might also be helpful to try using the observations without the use of the armbands to see if the same results would occur. It would even be useful to try the classification system with a different part of the operations (i.e. arrivals or on loads) or at a different station at a different airport⁵ to get a more representative opinion sample from the organization.

It is also recommended that questions be added to the survey to cover the topic of safety issues brought up in the pre-work meetings. It is unknown whether the information gathered regarding this subject during the Safety Perception Survey is accurate. The increase of safety issues in the pre-work meetings noted in the current study may have been caused by the presence of the observers. If this information is collected during a training classification system trial, this theory may be proved or disproved. Overall, the findings of the current study present a justifiable argument that the armband training classification program would be successfully accepted if implemented at *Company X*. It is speculated further that this system would work in any similar aerospace High Reliability/Low Commitment organization.

⁵ *Company X* conducts operations worldwide and stations exist at numerous locations around the planet.

REFERENCES

- Asch, S.E. (1995). Opinions and social pressure. *Scientific American*, 31, 193.
- Connor, M. (2001, June 28). Corrected update 1- *Company X* quarterly profits drop 54 percent. Retrieved on July 2, 2001 from World Wide Web: Address not included to maintain *Company X*'s anonymity.
- Endler, N.S. & Hartley, S. (1973). Relative competence, reinforcement and conformity. *European Journal of Social Psychology*, 3, 63.
- Fought, C. (2000). [Safety metrics package]. Unpublished software.
- Freiberg, K. & Freiburg, J. (1996). *Nuts! Southwest Airline's crazy recipe for business and personal success*. Austin, TX: Bard Press.
- Geller, S.E. (2001). *The psychology of safety handbook* (2nd. Ed). Boca Raton, FL: Lewis Publishers.
- Hess, T., Matteson, M., McNamara, K., & Thompson, R. (2000). [Safety perception survey at *Company X*]. Unpublished raw data.
- Koepnick, W. (1993, October). Do safety incentive programs really work? Presented at the National Safety Council Congress and Exposition, Chicago, IL.
- Krause, T.R. (1997). *The behavior-based safety processes: managing involvement for an injury-free culture* (2nd ed.). New York, NY; John Wiley & Sons.
- Lee, K.G. (2001). [Safety observation metric]. Unpublished software.
- McSween, T.E. (1995). *The values-based safety processes, improving your safety culture with a behavioral approach*. New York, NY; John Wiley & Sons.
- Milgram, S. (1963). Behavioral studies of obedience. *Journal of Abnormal Psychology*, 67, 371.
- Muchinsky, P.M. (1983). *Psychology applied to work: an introduction to industrial and organizational psychology*. Homewood, IL; The Dorsey Press.
- Paige, C.W. (1998). Flight Deck Operations. Retrieved on June 19, 2001 from the World Wide Web: <http://members.nbci.com/XMCM/charlespage/17FLIGHTOPS.htm>
- Reason, J. (2000). *Managing the risks of organizational accidents*. Aldershot, England: Ashgate Publishing Limited.
- Sekran, U. (2000). *Research methods for business* (3rd ed.). New York, NY: John Wiley & Sons.

APPENDIX A

Observation Raw Data

Safety Infraction	Date							Totals
	16-May	22-May	24-May	29-May	31-May	4-Jun	12-Jun	
Hearing protection	18	7	10	8	6	0	4	53
Improper lifting	10	6	5	6	3	0	0	30
Walking in between dollies when driver in tug	3	4	4	4	0	0	0	15
No honk before moving vehicle	5	5	2	0	0	0	0	12
Reckless behavior	0	2	0	0	1	0	5	8
No back belt used	0	2	0	2	0	0	0	4
Equipment not chocked properly	1	0	0	0	0	0	0	1
Cargo door not fully open before moving ladder	0	0	0	1	0	0	0	1
Dangerous goods mishandled	0	0	0	1	0	0	0	1
Cargo area not inspected	0	0	0	1	0	0	0	1
No safety net in place	0	0	0	0	0	1	0	1
Totals	37	26	21	23	10	1	9	127