
RESEARCH ARTICLES

Exploratory Analyses of the Effects of Managerial Support and Feedback Consequences on Behavioral Safety Maintenance

M. Dominic Cooper

ABSTRACT. Reviews indicate management commitment is vital to maintain behavioral safety processes. Similarly, the impact of observation frequency on safety behaviors is thought to be important. An employee-driven process which encompassed behavioral observations, goal-setting, and feedback was implemented in a paper mill with 55 workgroups using a within-group, time-series design. Data collected over 70 weeks show safety improved by 45% points with a corresponding 45% reduction of injuries. Management's demonstrable support was significantly associated with behavioral safety performance. Observation

M. Dominic Cooper is affiliated with School of Applied Health Sciences, Indiana University, Bloomington.

Address correspondence to: M. Dominic Cooper, School of Applied Health Sciences, HPER 116, Indiana University, Bloomington, IN 47405 (E-mail: domcoope@indiana.edu).

Journal of Organizational Behavior Management, Vol. 26(3) 2006

Available online at <http://jobm.haworthpress.com>

© 2006 by The Haworth Press, Inc. All rights reserved.

doi:10.1300/J075v26n03_01

1

frequency appeared to be unrelated to the observed percent safe scores. doi:10.1300/J075v26n03_01[Article copies available for a fee from The Haworth Document Delivery Service: 1-800-HAWORTH. E-mail address: <docdelivery@haworthpress.com> Website: <<http://www.HaworthPress.com>> © 2006 by The Haworth Press, Inc. All rights reserved.]

KEYWORDS. Behavioral safety, management commitment, observation frequency, goal-setting, feedback

Behavioral safety systems attempt to identify the small proportion of unsafe behaviors implicated in the majority of injuries. Once found, an examination of reinforcing contingencies identifies those antecedents and consequences supporting safe behavior. The safe behaviors and/or the results of safe behaviors (e.g., creating safe work conditions) are placed on workgroup specific checklists. Typically, trained observers make use of these to monitor and record the safety behavior of their peers on a regular basis (e.g., daily or weekly). Recorded observations are computed to provide a percentage safe score. This forms the basis for feedback via graphic charts or written performance summaries, so individuals or workgroups can track their progress against self-set or assigned improvement goals (Locke & Latham, 1990). Observers may also provide verbal feedback at the point of observation to facilitate corrective actions (Algera, 1990; McAfee & Winn, 1989). Research indicates 50-75% reductions in injury rates are possible within a relatively short period (Cooper et al., 1994; Laitinen, Marjamaki, & Paivarinta, 1999) with the impact lasting for many years (Fox, Hopkins, & Anger, 1987).

Both reviews (e.g., Grindle, Dickinson, & Boettcher, 2000) and commentaries (Austin, 2000) indicate the need to empirically examine the factors leading to long-term program maintenance and performance change. Studies evaluating the functional components of behavioral safety have tended to examine the impact various combinations of particular antecedents (e.g., training and goals) and consequences (e.g., feedback, praise, and incentives) exert on performance. The effects of management's support on the observed percentage safe score and/or Observation Frequency are often discussed but rarely empirically vali-

dated in the literature. It is expected that both of these are essential to the maintenance of any behavioral safety intervention.

MANAGEMENT COMMITMENT

In the current paper, management commitment is defined as “engaging in and maintaining behaviors that help subordinates achieve a goal.” High levels of management commitment are essential for organizational interventions to succeed (Rodgers, Hunter, & Rogers, 1993). Safety interventions are no exception. Work conducted in the UK construction industry revealed management’s commitment to the behavioral safety process accounted for approximately 31% of the variance in the observed percent safe score (Robertson et al., 1999). In an Australian qualitative study of multiple-site interventions, Harper et al. (1997) identified nine important leadership features essential to the maintenance of behavioral safety initiatives. Active managerial involvement in the behavioral safety process was the most important. Likewise, in a Canadian study within the petrochemical industry, Cook and McSween (2000) indicated those behavioral safety processes with high supervisory involvement also had a higher percentage of employees conducting observations (i.e., observation frequency).

Broadly speaking, management commitment is measured in two ways: Direct questions are asked of managers (Hollenbeck et al., 1989), or their commitment behaviors are monitored (Salancik, 1977). The difficulty with the former method is managers seem unlikely to express, or admit to, the notion that they are not committed to safety or the initiative. The advantage of the latter method is that proof of commitment for the program are the same management behaviors that promote the program goals, and these behaviors can be reinforced like any other operant.

Some causes of failed safety initiatives can be found in the behavior of the leadership team and the signals which they communicate to the workforce (Buchan et al., 1999). Zohar (2000) utilized a behavioral leadership-based intervention model to show improved transactional behaviors of supervisors resulted in greater employee use of hearing protection,

and significantly decreased the number of minor injuries. Zohar and Luria (2003) obtained similar results in an oil refinery and two food processing factories. Zohar suggested the influence of supervisory behavior on employee behavior becomes stronger due to successive reinforcement-learning effects derived from supervisory provision of antecedents (e.g., goals) and consequences (e.g., feedback). In addition, the reinforcement value of personal supervisory attention and recognition is thought to play a large part in changing employee behavior (Stajkovic & Luthans, 1997).

Engaging in such behaviors may also become naturally, and perhaps vicariously, reinforcing for the managers and supervisors themselves as the desired outcomes are realized (Bandura, 1986). Thus, a reciprocal relationship appears to exist between managerial commitment behaviors and employee safety behaviors, which lead to increase in both over a period of time.

Managerial commitment behaviors can be wide and varied. However, they are likely to reflect both people issues (e.g., manager speaking to an employee to correct an unsafe act) and system support issues (e.g., manager monitoring project progress). One pertinent system support factor for all behavioral safety processes is observation frequency. This is an important indicator of the company's commitment to the behavioral safety process because managers must choose to incur the financial costs involved in releasing "observers" from their place of work for approximately 15-30 minutes at a time. For example, some behavioral safety processes are based on one observation per work area, per week (e.g., Mattila & Hyodynmaa, 1988), while others are conducted in all work areas 2-4 days per week (e.g., Reber, Wallin, & Chokkar, 1990) or daily (e.g., Cooper et al., 1994). Often, observations do not take place at all due to line-managers' resistance to the process (Cooper, 1998; Krause, 1995).

OBSERVATION FREQUENCY

Observation frequency is thought to be important in behavioral safety processes for a number of reasons: (1) a large number of observations create more reliable (Bausell, 1986) percentage safe scores; (2) larger

numbers of observations make it easier to account for variations in performance (e.g., Komaki, Zlotnick, & Jensen, 1986); (3) a greater number of observations provides greater potential for more frequent verbal feedback at the point of observation; (4) the very act of observing is thought to change the behavior of the observer (Alvero & Austin, 2004) and the observed (Kazdin, 1982); and (5) conducting observations provides an indirect measurement metric of an observer's commitment to the process (Robertson et al., 1999).

Persuading employees to conduct behavioral observations is fraught with difficulty. Often perceived as "spying," some US labor unions (e.g., UAW, TWU) officially disapprove of behavioral safety as they believe behavioral safety systems: (1) lead management to abrogate their responsibility for safety; (2) treat employees as "Pavlovian" dogs to be manipulated; (3) generate fear amongst workers; (4) create conflict among workers; (5) lead to less use of higher engineering controls; and (6) discourage the reporting of injuries and drive problems underground (Frederick & Lessin, 2000). Some of these assertions have been rebutted (Cooper, 2003), but they remain very important issues awaiting scientific inquiry. They are also very real obstacles to overcome in the workplace when introducing a behavioral safety system.

In sum, the above evidence indicates management's commitment and Observation Frequency are both important variables helping to determine the effectiveness and potential maintenance of a behavioral safety process. The current study aimed to explore the relationship between the percentage safe score, the Observation Frequency and management's commitment in a paper mill.

METHOD

Participants and Setting

This study was conducted in a paper manufacturing plant employing approximately 500 unionized personnel within nine main departments.

Two main production departments (Paper Mills One and Two) operated with shift employees who worked on a continuous 24-hour, 3×8 hour rotating shift system with two shifts off-duty at any one time. Within these two departments, there were seven main work areas. Employees directed diluted wood pulp onto a paper machine wire (about 9 meters wide, moving at a speed of 45 mph) at a rate of 2 tons per second. They, then, ensured the newly formed paper web left this section and was picked up by a vacuum and forwarded to be pressed between a series of rollers (weighing over 60 tons each) to eliminate water. Subsequently, the paper web was processed through a steam-heated “dryer” section and “pressed” between polished steel rollers. Employees had to react if the paper web breaks at this point. Thereafter, the paper web was scanned for weight, thickness, and moisture, before being reeled onto a jumbo reel. A jumbo reel contained almost 25 tons of paper, and was changed every 50 minutes. After being changed, the completed jumbo reels were then forwarded to the “slitting area” where they were cut to size to produce 30 customer reels using winding machines that run at speeds of up to 90 mph. Once cut, the reels were wrapped and taken to the warehouse with forklift trucks.

Employees within warehousing operations worked a 2×8 hour shift system (morning and afternoon). The remaining personnel were 18 workgroups, comprised office workers, laboratory staff, engineering maintenance, and contractors (e.g., canteen, cleaners) who worked a “normal” 5-day, 39-hour week. During shutdowns for maintenance there could also be up to 5,000 extra contractors on-site for a week or more at a time.

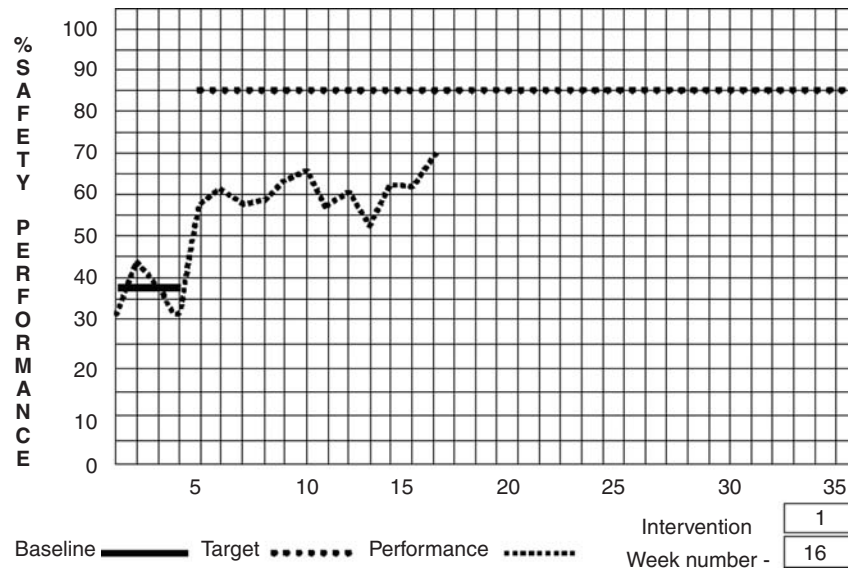
The company had one full-time safety professional and one management-union safety committee in place prior to implementation. In addition to hazard spotting exercises (i.e., line-managers and workers looking for unsafe conditions), employees were encouraged to report all potential incidents, no matter how minor. Initial analyses of these suggested a large number of unsafe behaviors were taking place, which led the site to explore the possibility of implementing a behavioral safety process. The site drew up a working group drawn from a cross-section of the mill and all the union safety representatives. The working group visited

many other companies implementing different behavioral safety processes and spoke with many safety experts to select a system compatible with the company's culture and aims.

Intervention Design

This quasi-experimental case study used a within-group, time-series design (Komaki & Goltz, 2001) with 55 workgroups which included a transitional workforce of 50-150 on-site contractors. While data were collected continuously over 70 weeks, three consecutive interventions were implemented using an ABABAB design. The sequence of interventions included: (1) Baseline 1 (4 weeks); (2) Intervention 1 (21 weeks); (3) Return to Baseline 2 (4 weeks); (4) Replication Intervention (18 weeks); (5) Return to Baseline 3 (4 weeks); and (6) Replication Intervention (19 weeks). Each baseline started immediately after the preceding intervention (see Figure 1).

FIGURE 1. Example Graphic Feedback Chart



Measures

Safety Performance. An examination of injury incident records for the previous two years helped to identify the most frequent types of unsafe behaviors. Unsafe behaviors were initially categorized according to their geographical location in the mill, the types of injuries sustained, and the place of injury on the body. Then each was subjected to a functional analysis conducted by previously trained employees ($n = 30$) who had specifically volunteered to help the project coordinators develop the behavioral checklists after their work shift. Teams of five volunteers collectively performed each functional analysis.

An example unsafe behavior causing injury was “Striking bar with 14lb sledgehammer to remove wad between drums, while someone is under winder drums in the pit below.” A prior injury occurred when vibration from this action caused the employee to release the bar which fell on a colleague below and struck him on the face. The identified antecedents for this example included the lack of a proper tool for removing wads, the availability of improvised tools, and the end of the shift approaching (i.e., a time constraint). The strongest consequences maintaining this unsafe behavior were saving time, getting the job done, and keeping production going. In addition to identifying the appropriate safe behavior, these analyses also revealed barriers to safe behavior such as, in this example, the lack of a formal safe system of work, proper tools for the job, and availability of improvised tools (i.e., sledgehammer).

For each of the 31 observation zones across 21 different work areas, unsafe behaviors identified by the functional analysis teams were listed on draft checklists and returned to employees in every work group for comment at special meetings. Each checklist contained a number of proposed safe behaviors, the reasons for their inclusion (e.g., a high number of injuries on record due to the behaviors), and a rating scale of “work area agreement” from A to E for each proposed checklist item (A being “Strongly retain” and E being “Strongly eliminate”). In some instances, the workgroups removed or added behaviors. Before final adoption, the workgroup approved each behavior on the checklist.

Each adopted checklist (see Table 1 for a sample checklist) contained a maximum of 20 safe behaviors (e.g., personnel are holding the handrails when using stairs) or “footprint conditions” (e.g., hoses are coiled and stored on hooks, not left in walkways) pertaining to the work area of interest. Footprint conditions were defined as “a physical trace that an unsafe behavior has occurred.” For example, an uncoiled hose left across a walkway to cause a trip hazard provided evidence that someone has not put the hose away after use (an unsafe behavior). Such an event was referred to as a “Footprint condition” simply because the hose could not have gotten there on its own (others call these “results”; see Daniels, 1989). Therefore, a “footprint” of an unsafe behavior usually created an unsafe “condition.” These behaviors and footprints were divided into various categories (e.g., Housekeeping, Personal Protective Equipment).

Each checklist contained three columns: Safe, Unsafe, and Unseen (see Table 1) for observers to record the results of their observations of employee safety behavior (see procedural section below for observation details). If any one person in a workgroup performed an unsafe behavior, the whole workgroup was scored unsafe on that particular behavior (Komaki, Barwick, & Scott, 1978). Therefore, any particular behavior recorded as *safe* meant the whole workgroup was successfully performing a safe behavior. A frequency count of the number of persons unsafely performing a particular behavior determined the recording of unsafe behavior. The total number of safe behaviors recorded were divided by the sum of the total safe and unsafe behaviors recorded, and multiplied by 100 to calculate an Observed Percent Safe score (the primary dependent variable in this study).

The unseen column was marked when a particular behavior or footprint did not occur during the 15-30 minutes of observation tour (the project team analyzed these to remove infrequently recorded behaviors from subsequent intervention checklists).

Project coordinators entered daily observation data into a behavioral safety computer database (Cooper, Bissett, & Walters, 1999) when they received the completed behavioral checklists from the workgroup observers. The database contained an exact copy of each checklist, for each observation zone, by shift or workgroup, with corresponding data

TABLE 1. Example Behavioral Safety Checklist

Company: Paper Mill		Department: Paper Production		Work Area: PM1A		
Intervention: 1		Team: A		Date:		
Category: Access		Safe	Unsafe	Unseen		
1.0	Walkways are clear of obstructions					
Category: Housekeeping/General		Safe	Unsafe	Unseen		
2.0	All guards are in place and in good condition					
3.0	Personnel are not standing on handrails					
4.0	Back is kept straight when lifting					
5.0	Crane hooks are left at safe height					
6.0	Floors are clear of spillage's of:					
6.1	Oil					
6.2	Chemicals					
7.0	Sharp blades are in containers when not in use					
8.0	All drain grids are in good condition and seated properly					
Category: Personal Protective Equipment (PPE)		Safe	Unsafe	Unseen		
9.0	Ear protection is worn at all times on operating floor					
10.0	Full face mask is being worn when tail feeding on calendar stack					
11.0	Gloves are worn when handling:					
11.1	Doctor Blades					
11.2	Hot objects					
11.3	Chemicals					
Category: Emergency Equipment		Safe	Unsafe	Unseen		
12.0	Emergency exits are free from obstruction					
13.0	Red fire hoses are used for emergency purposes only					
14.0	Emergency equipment is free from obstruction					
15.0	Fire hose boxes all contain the full & correct # of items					
TOTAL:						
%Safe = (total safe/(total safe + total unsafe)) * 100						
Comments:						

entry fields in the safe, unsafe, and unseen columns. Once entered, the program automatically calculated a percent safe score (i.e., total safe/total safe and unsafe, multiplied by 100). The program was used to generate the weekly feedback reports for each workgroup.

Observation Frequency. This study was designed so that 251 observations could potentially be conducted per week across all the observation areas and shift workgroups. Observation Frequency (OF) was calculated by dividing the number of *actual* observations recorded by the *potential* number of observations (i.e., 251), and multiplying by 100. Missing observations were tracked using the computerized behavioral safety database, making it possible to accurately record the number of *actual* observations completed (an observer's completed checklist entered into the program counted as one completed observation). The coordinator entered the shift rotation or work pattern for each of the observation zones within the work areas into the computer program at the beginning of the study, and updated the shift rotation pattern annually. This provided the expected number of observations for each workgroup in the observation zones according to shift patterns.

In the current paper, OF is expressed as the aggregated average weekly percentage of observations conducted across the site as a whole. Observation Frequency was monitored throughout the study starting at the beginning of the baseline period.

Management Commitment. Briefing or "buy-in" sessions (described later in the Procedure) took place at the beginning of the project. The whole management team was asked to "pledge" their support to the initiative by agreeing to engage in specific behavioral acts to demonstrate their commitment. These management commitment behaviors were monitored by using a behavioral Visible Ongoing Support (VOS) checklist containing the seven items depicted in Table 2.

Each set of workgroup observers were asked to complete the VOS measure once a week, on the basis of their recollection of support received from the various parties (i.e., line-manager, project coordinator, etc.), or activities undertaken (e.g., feedback meetings held). The VOS checklists contained two columns: Yes and No, which were used to calculate the percent VOS in the same manner as the percent safe

TABLE 2. Visible Ongoing Support Behaviors

Number	Item
1	Has any member of management accompanied you during an observation this week?
2	Has your champion made contact to support you in your role, this week?
3	Has the process coordinator made contact to support you in your role, this week?
4	Has a weekly feedback meeting been held this week?
4a	If so, were the previous week's findings discussed?
4b	Were everybody's questions answered satisfactorily?
5	Have any remedial actions been addressed since last week?

score for each individual observation zone. The VOS ratings began immediately after the baseline period was completed because the managerial commitment behaviors were not required until then.

Remedial Action Rate. In this study, a remedial action was defined as “any action that the organization took to rectify a physical hazard that has the potential to cause harm, which could not immediately be resolved at the time of identification by an individual without recourse to a formal organizational request procedure.” The company strongly encouraged employees to report hazards so they could be dealt with before an accident occurred. Any hazards spotted by an observer that could not immediately be dealt with and required a potential remedial action (e.g., the painting of steps to the medical facility in yellow to highlight their potential as a tripping hazard) were noted in a specific comments section placed at the bottom of every behavioral observation checklist.

These were recorded in the computer program by the project coordinator at the time of observation data entry and assigned a special code to make these distinguishable from those reported by all employees. The coordinator then directed these to the appropriate authority for action (e.g., engineering department).

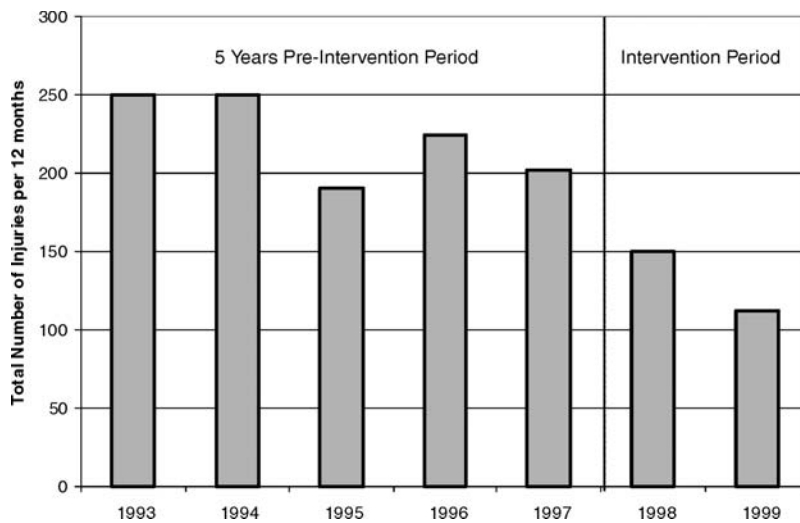
The weekly remedial action completion index was calculated by dividing the number of completed actions by the total number of hazards identified, and multiplying by 100. In an attempt to maintain workforce enthusiasm for, and commitment to, the project, examples of the items

addressed were also published in a monthly newsletter to publicize the tangible results the initiative was producing.

Injuries. The company involved in the study had a policy requiring employees to report all injuries regardless of the need for medical treatment. Injury was defined as “damage or harm done to or suffered by a person.” Employees reported these via the company’s official “accident reporting book.” The company was required by law to report serious injuries (e.g., temporary loss of sight, unconsciousness) immediately to the British Health and Safety Executive. In addition, all injuries requiring three or more days away (UK standard for a lost-time injury) from normal duties had to be reported to the regulator within 10 days.

Whether or not all injuries were actually reported is a moot point. Figure 2 shows the historical figures for all known injuries (lost-time and minor) for the five years prior to implementation (January to December). In the year prior to the beginning of the behavioral safety process, the company had experienced 202 injuries. The number of injuries was the primary outcome measure used by the company to assess the effectiveness of the behavioral safety process.

FIGURE 2. Total Number of Injuries Pre- and Post-Interventions



Procedure

Overview of Intervention 1

The implementation procedure steps followed for Intervention 1 were: (1) Obtaining employee agreement and participation; (2) recruiting and training the project team; (3) developing checklists; (4) recruiting and training employee observers; (5) establishing work group baselines; (6) setting improvement goals; and (7) monitoring performance and providing feedback. The replication interventions repeated the procedural steps above but did not repeat the employee agreement and project team training sessions. Therefore, these later interventions started with the development of new checklists and training new employee observers.

Obtaining Employee Agreement and Participation. To facilitate workforce acceptance, the author conducted initial briefings with *all* site personnel, including contractors, at one-hour meetings *prior* to the recruitment, training, and implementation process. Personnel were informed about procedural steps involved in the behavioral safety process and what it meant to each of them in terms of (1) helping to develop behavioral checklists; (2) being observed or observing; (3) setting behavioral improvement targets; and (4) being provided with detailed feedback about their own workgroup's weekly safety performance.

Personnel were then asked whether they would participate in the implementation of such a system by being part of the project team, an observer, or by allowing themselves to be observed. One issue that surfaced at every briefing was company discipline resulting from observations of unsafe behavior. The company's most senior management attended the meeting and provided oral assurances that discipline for non-compliance with the behaviors on the observation checklists was not an option. However, management did retain the right to discipline people who clearly and knowingly put other people at risk of injury or death.

Project Team Training. A project team comprising two employee coordinators (i.e., paid by the hour) were recruited along with two safety champions drawn from the senior management team. Over a five-day period the project team, other members of the senior management

team, and the site's Trades Union Council were trained in the principles and strategies of implementing a behavioral safety system. In essence, this interactive training covered a six-stage process encompassing: (1) Functional analysis of accident and incident records; (2) development of behavioral observation checklists; (3) observer training; (4) establishing a baseline; (5) participative goal-setting (Cooper et al., 1992); and (6) feedback mechanisms.

Developing Checklists. Immediately after the project team training, 30 additional volunteer employees were trained on a one-day course in the principles of functional analysis to help assist the project team in the development of the behavioral checklists. The volunteers were those who had indicated their willingness to become involved with the project's development at the initial briefings. Once the checklist development was completed, the observer recruitment began.

Observer Recruiting and Training. The project team was then tasked with observer recruitment. Three criteria were used: (1) Observers needed to be committed to improving safety; (2) observers needed to be willing to observe their colleagues on a daily basis for approximately six months (an intervention period); and (3) the project team had to recruit an observer from every workgroup.

Initially, the union safety representatives were approached to be observers. All agreed to observe and help recruit other observers. Notices and memos were also posted around the site asking for volunteer observers. The CEO signed and distributed a statement from the senior management team, entitled "Nothing we do here is worth getting hurt for," and offered the management team's full support.

Recent remedial actions that addressed previously unsafe conditions were also publicized to demonstrate to employees that the company was serious about improving safety. Examples included modifying a conveyor to eliminate a head hazard; painting various tripping hazards in yellow (i.e., steps and bund walls to highlight them from the background); designing and building hosepipe holders.

Volunteer observers (n = 195) were recruited and trained within 12 weeks of the initial briefings. Observer training was as an open event so

that employees who were skeptical could attend. Attendees were not required to act as observers.

Every observer attended a one-day eight-hour training course to (1) learn how to observe using the behavioral checklists; (2) how to give verbal feedback at the point of observation and at weekly feedback meetings; and (3) how to set participative goals with their workgroups. Observers also conducted practice “on-site” observations with their respective workgroup checklists.

Over a subsequent two-week period, observers conducted 15-30 minutes practice observations every day they were at work (i.e., for two shift cycles) but the data were not used to formally compute a workgroup’s percent safe score. The coordinators also conducted checks to ensure observers were using the scoring system correctly and they were comfortable in their duties. Comparisons of shift observations identified any misinterpretations of the behavioral items between observers. For example, if A and C shift observers indicated housekeeping in a certain area was poor, but B shift observer indicated it was very tidy in the time between A and C shift observations, the coordinators checked the physical housekeeping status, and coach observer B on interpretation and scoring if warranted.

It is important to note the coordinators did not collect any formal inter-observer agreement data throughout the study. Their sole concern was to get employees to observe in as accurate a manner as possible. One way of helping to ensure ongoing observer accuracy was to ask line-managers to accompany an observer at least once per week to conduct a paired observation.

Establishing Baselines. After observation training and practice, a four-week observation period was used to establish a baseline before setting an improvement goal. Employees did not receive training in the behaviors on the checklists (each employee would eventually go on an industry-specific three-day, generic working safely course after the study). Line-managers informed employees at the beginning of the baseline period that observers would begin to observe once per day, at random times, for 15-30 minutes. At this time, observers posted the work area’s checklist on a notice board. This served the dual practical

purpose of reminding personnel of the behaviors being monitored in that area and minimizing the probability of observers being accused of spying. Given the assistance of the workgroups in developing the observation checklists, there may have been a small positive effect on safety performance during the baseline period (Alvero & Austin, 2004). If so, the result would be a slightly inflated baseline score making the study results more conservative than they may otherwise have been.

To establish a baseline for a work area, each observer monitored everyone in his or her workgroup once per day for 15-30 minutes over four weeks (or four shift cycles). Each observer chose the time of day during a work shift when his or her observation would occur. Thus, the time of day observations took place varied. Observers did not give verbal feedback during this period, unless someone was in imminent danger of hurting themselves or their colleagues. During the baseline period, there was no formal written or posted graphic feedback about safety performance.

Goal-Setting. At the end of each four-week baseline period, each workgroup and senior managers attended their respective work area goal-setting meeting. All workgroup members participated in setting a goal (see Cooper, 1993). To set a goal, employees were first shown how the percent safe score was calculated. They were then informed that the baseline average percent safe score was the aggregate mean of all the baseline observations for all the behaviors on that workgroup's behavioral checklist. The group was then informed of the percent safe baseline average for the previous four weeks (inclusive of the five best and worst scoring behaviors). This was presented in graphic form on that workgroup's graphical feedback chart, and in a written summary. Employees were encouraged to discuss why they thought they had done well on some behaviors and not so well on others. Those managers in attendance also took notes on actions required to help improve the situational aspects that led to some of the unsafe behaviors (e.g., the right equipment not being in the right place at the right time).

After these discussions, every employee was asked to publicly indicate how much more safely (in percent safe terms) he or she thought the group could work over the intervention period. A goal was not set for each target behavior, but for all the behaviors on the checklist, as a

collective “whole.” The workgroup observer recorded each individual’s suggested goal level or “pledge.” These were summed and divided by the number of employees present to calculate a suggested workgroup’s average percent safe target level. Each employee was then asked if he or she agreed with this average as the workgroup’s goal. Observers posted their workgroup’s agreed upon goal on a dedicated graphic feedback chart as a horizontal line at the appropriate percent safe goal level.

Monitoring. For 21 weeks after the goal-setting meeting, each work area observer continued to monitor his or her colleague’s safety behavior on a daily basis for 15-30 minutes. Again, each observer chose the time of day during a work shift when his or her observation would take place. In principle, Visible Ongoing Support (VOS) was also monitored once a week by the observers beginning after each workgroup had set its safety improvement goal.

Feedback. Observers were encouraged to give verbal feedback at the point of observation. In addition, a computerized behavioral safety-tracking program (Cooper et al., 1999) analyzed the observation scores for each workgroup on a weekly basis. Each observer would hand in his or her completed behavioral checklist to a project coordinator for data entry (usually at the end of the day/shift on which the observation was conducted). The weekly percent safe score for each workgroup was calculated and posted on their graphic feedback chart. These had been placed on walls in the respective working areas (see Figure 1). This represented the group’s average score for the week calculated from the observation totals entered each day.

The tracking program also produced a written report of results by category of behavior (e.g., Use of tools, Housekeeping), and highlighted the five highest and lowest scoring percent safety behaviors for that week (see Table 3 for an example of the report). Each workgroup attended a weekly 30-minute meeting to discuss the workgroup’s ongoing safety performance presented on the written feedback form. Each person was given a copy, but it was not publicly posted on work area notice boards. The written feedback was deemed ‘confidential’ to each individual workgroup. Neither did management or employees receive explicit verbal, graphic or written feedback regarding VOS performance.

TABLE 3. Example Weekly Team Briefing Feedback Report

Company: Paper Mill	Department: Paper Production	Work Area: PM1 A
Intervention: 1	Week: 6	Start: 10/4/1998
		End: 10/10/1998
Team: A		
Target: 86.00	Weekly Average: 61.96	Previous Weeks Average: 58.00

Overall Average Safety Performance increased by: 3.96

Number of Observations Conducted: 7 **Actual Baseline Percentage: 36.51**

Summary of Checklist Categories:

Category Observations	Safe	Unsafe	Unseen	Percentage of Safe	Percentage of
Access	0	18	0	0.00	19.57
Housekeeping/General	34	6	16	85.00	43.48
PPE	2	5	28	28.57	7.61
Emergency Equipment	21	6	1	77.78	29.35
Grand Totals:	57	35	45	61.96	100.00

Best Scoring Checklist Items

Item	Checklist Item Text	Safes
8.0	All drain grids are in good condition and seated properly	7
14.0	Emergency equipment is free from obstruction	7
6.2	Floors are clear of spillages of Chemicals	6
2.0	Guards are in place and in good condition	6
12.0	Emergency exits are free from obstruction	6

Worst Scoring Checklist Items

Item	Checklist Item Text	Unsafe
1.0	Walkways are clear of obstructions	18
10.0	Ear protection is worn at all times on operating floor	5
13.0	Red fire hoses are used for emergency purposes only	4
6.1	Floors are clear of spillages of Oil	3
5.0	Crane hooks are left at safe height	2

Replication Interventions

In this study, there were three interventions with each workgroup. The two additional interventions repeated the above implementation process excluding the initial workforce briefings and project team training. Each new intervention used adapted behavioral checklists and new workgroup observers. The change process for these interventions included:

Checklists. Eight weeks prior to the start of a new intervention an analysis of each workgroup's observation data was conducted to ascertain (1) behaviors observed as 100% safe; (2) behaviors not frequently observed; and (3) behaviors still representing a problem. Behaviors not observed frequently or were 100% safe for an extended period (usually the last eight weeks or more) were replaced by other behaviors. Replacement behaviors were either extracted from an analysis of the company's potential incident reporting system, recommended by the workgroup, or noticed as problems by the observers during the previous intervention. Each workgroup approved their new behavioral checklist prior to its use.

Observers. One volunteer observer was again recruited from each workgroup ($n = 55$). All the observers were completely new observers. New observers practiced their observations alongside existing observers during the last week of the prior intervention. The practice data was used solely for the purposes of ensuring the new observers were competent in their task.

Baselines. Prior to the start of each intervention, a baseline was re-established for each workgroup so new workgroup specific improvement goals could be set. New baseline observations began immediately after the end of each intervention. Feedback ceased entirely (verbal, graphic, and weekly meetings) during the four-week baselines between each intervention. Workgroup improvement goals were not set until the week after the baseline period ended. Thus, no formal project-induced antecedents or consequences were in operation during these baseline periods even though the observations continued.

Intervention Handover. At the end of each intervention period, the final weekly feedback meeting highlighted the successes of the whole intervention. Current observers were thanked for their efforts, and the

new observers drawn from the workgroup were formally introduced along with the new behavioral checklist to be used over the next intervention period. The workgroups were reminded the next feedback meeting would be conducted after a break of five weeks so that a new baseline could be established with the new behavioral checklist in the intervening four-week period.

Goal-Setting and Feedback. Exactly the same process was used to determine each workgroup's goal. The graphic feedback charts were "cleared" of the previous interventions' results, and re-used to deliver the results for that specific intervention. Written feedback, highlighting the best and worst scoring behaviors, was delivered in the same way at weekly meetings led by the work group observers.

Independent Variable Integrity

To establish events occurred as planned, each of the independent variables was monitored formally or informally either via the computer tracking software by the project team or a combination of both.

Observation Frequency. Tracked via a computerized behavioral safety-tracking program (Cooper et al., 1999).

Goal-Setting. To ensure a goal was set for each workgroup, a project team member attended each of the work area goal-setting meetings. Every planned goal-setting session took place during each of the three intervention periods.

Feedback Meetings. Observers received a printed "weekly feedback report" specific to their shift group or work area on the day the feedback meetings were scheduled. To help ensure the feedback meetings actually took place, a member of the project team would, on a random basis, attend one or two each week throughout the projects duration. Weekly VOS data helped track the occurrence of such meetings. The author also conducted an initial audit of all available records, and interviewed employees. When these methods revealed that not all feedback meetings were taking place as intended, meetings with all supervision and management stressed the importance of feedback to improve behavioral performance. Regular informal follow-up audits by project team

members indicated approximately 90% of subsequent feedback meetings took place across all three interventions with no consistent lack of weekly meetings in any one particular work area.

Graphic Feedback Data. Observers did not receive computed percent safe observation data during any baseline period. Observers received a summary of their work area's weekly baseline scores and overall baseline average only on the day of their goal-setting session. Observers updated the graphic feedback charts at the weekly feedback meetings. The project coordinators and champion would check that these had been updated as intended, during their tours of the plant. All indications were that these were completed 100% of the time.

Verbal Feedback Data. During each intervention phase, the project team did not attempt to formally monitor the amount of verbal feedback provided by observers to those observed.

Visible Ongoing Support (VOS). Work area observers recorded VOS on a weekly basis once a baseline period was complete. Observers received a VOS checklist to indicate if support activities occurred (e.g., line-managers accompanying them during an observation) when they had given the written weekly feedback records.

RESULTS

Safety Performance

For the initial Intervention the four-week baseline score was 49.3% (range: 46.33-51.83; $SD = 2.27$). This indicates that half of the behaviors observed were performed safely when the project began. The site's average improvement goal was 70%. The mean average percent safe score (70.6%) suggests this intervention goal was achieved. Figure 3 reveals an upward trend in safety behavior for the first nine weeks of the initial intervention suggesting goal achievement after six weeks. Employees maintained this level of performance for the remaining 15 weeks of the intervention. The annual winter holiday break (Weeks 15

FIGURE 3. Average Weekly Percent Safe Scores For Each Intervention



and 16) and a large two-week shutdown (Weeks 21-22) using external contractors may account for the two dips in the time-series analysis.

The average baseline score before the second intervention was 68% (range: 66-69; *SD* = 1.41), which indicates overall safety behavior improved 18.7% points over the original baseline. The site’s improvement goal was set at 77% for this intervention. The mean average percent safe score during the second intervention was 77.9% (range: 62-94; *SD* = 7.96). Again, employees achieved the goal within six weeks. They maintained or exceeded this for the remainder of the intervention (12 weeks).

The average baseline score before the third intervention was 75.75% (range: 74-80%; *SD* = 2.87), which was a 7% point improvement compared to the prior baseline, and a 25.7% point improvement compared to the original baseline. The site’s improvement goal was set at 85% for this intervention. This goal was attained within 11 weeks of the intervention onset and employees maintained this level of performance for the remaining 8 weeks of the intervention. However, the overall mean average percent safe score for the third intervention was 83.7% (range:

73-94%; $SD = 7.33$) which fell slightly short of the 85% site intervention goal. In part, this was due to another maintenance shutdown in week 53 where an additional 500 contractors were on site and the winter holiday break in weeks 65-66.

When compiled in four weekly averages to coincide with the equivalent monthly injury numbers, a small non-significant negative correlation ($r = -0.016$) was obtained between average percent safe scores and the number of injuries over the study period.

Observation Frequency (OF)

In total, employees conducted 14,146 safety observations throughout the study period. Employees conducted 5,506 daily observations during the first intervention, 4,185 during the second intervention, and 4,455 during the third.

The average Observation Frequency was 87.75% (range: 82-93; $SD = 2.89$) during the first intervention, 75.8% (range: 72-82; $SD = 3.38$) for the second intervention, and 77.18% (range: 82-93; $SD = 2.97$) for the third intervention. These results suggest observers in the latter two interventions were not as committed to the project as those observers of the first intervention. Figure 4 shows the weekly observation frequency for each intervention. A comparison of behavioral safety performance throughout all three interventions does not reveal any apparent correlations with OF. Similarly, although in the right direction, a small negative, but non-significant correlation was obtained between Observation Frequency and number of injuries ($r = -0.226$, n.s.) when compiled in four weekly averages to coincide with the equivalent monthly injury numbers.

Visible Ongoing Support (VOS)

Observers recorded VOS 786 times, from a possible 3,190 (24.6%). Observers completed 288 VOS recordings during the first intervention, 253 in the second, and 245 in the third. Weekly VOS scores were aggregated to compose average VOS scores for each intervention. During the

FIGURE 4. Weekly Average Percent Observation Frequency Rate Per Intervention



first intervention, 25% (range: 10.81-46; $SD = 10.33$) of scheduled weekly VOS recordings were completed. During the second intervention, 26% (range: 10.8-48.11; $SD = 11.00$) were completed. During the third intervention, an average 23% (range: 12-45.26; $SD = 9.32$) were completed.

The managerial Visible Ongoing Support (VOS) rated by observers averaged 56.1%, (range: 32.5-75; $SD = 11.33$) over the 21 weeks of the first intervention. The second intervention's VOS scores averaged 70.6%, (range: 59-77; $SD = 5.96$). The third intervention's VOS scores averaged 77% (range: 72-81; $SD = 5.08$). Thus, mean VOS levels increased over each intervention (see Figure 5).

The VOS scores were also compiled in four weekly averages to coincide with the equivalent monthly injury numbers. A small, negative, non-significant correlation ($r = -0.28$) was obtained between these monthly VOS average percent scores and the number of injuries over the 70 weeks.

To identify which specific VOS behaviors were routinely completed, calculations were made of each behavior during the intervention phases

(see Table 4). This revealed that the project coordinators' support increased across the three interventions, as the champions (i.e., senior managers who championed the behavioral safety process throughout the company) decreased. Interestingly, line-managers accompanied observers much more in the second intervention, when VOS levels

FIGURE 5. Weekly Average Percent Visible Ongoing Support Scores Per Intervention

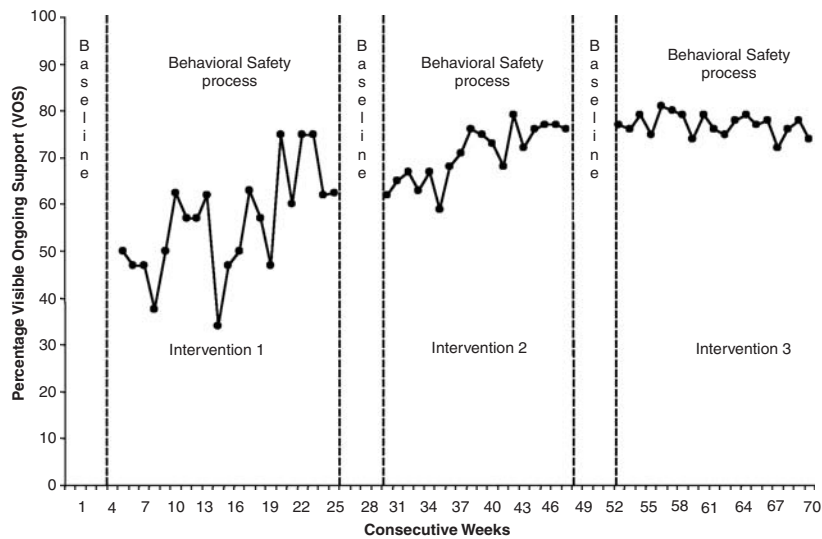


TABLE 4. Percentage of Visible Ongoing Support Behaviors Recorded by Intervention

Item #	Text Description	Int. 1	Int. 2	Int.3
1	Accompanied by a manager	6	16	4
2	Champion supported observer	13	9	1
3	Project coordinator supported observer	15	24	27
4	Held a weekly feedback meeting	21	15	20
4a	Discussed previous week's results	20	13	18
4b	Everyone's questions answered satisfactorily	15	11	16
5	Remedial action closed since last week	10	12	13

tended to increase. The number of weekly feedback meetings appears to decline in the second intervention, but this may be related to the increased managerial and coordinator support of observers. Most feedback meetings discussed the previous week's results with everyone's questions answered satisfactorily. Finally, the completed remedial actions appeared to be relatively constant across the interventions.

Remedial Action Rate

The workforce reported 216 physical hazards. Of these, 151 remedial actions were completed during the project period representing a 70% completion rate ($\# \text{ completed items} / \# \text{ reported items} * 100$). However, the researcher did not receive these data in a weekly format, making it difficult to assess its impact on the recorded percentage safe score or the impact such actions had on employee buy-in to a behavioral safety process.

Injury Reduction

Compared to 202 injuries in the 12 months prior to the study (see Figure 2), the total number of injuries reduced to 150 in the first 12 months of the intervention, representing a 24.75% reduction. In the following 12 months, total injuries decreased to 112, which was a further 25% reduction (the observation data reported in this study reflects only the following five months, but the site continued to maintain the process). It would appear, therefore, in conjunction with other industry-led initiatives, and the increased emphasis on safety in general, the project helped to reduce the site's total number of injuries by about 45%.

Cost/Benefits

The costs of the project included labor hour costs for safety meetings, training, and observations plus a 25% overhead for statutory employer contributions and the author's professional fees. Labor costs largely

refer to project related “off-task” costs whereby personnel were not involved in production when they were involved in the project. Salary costs for the one full-time (£33K) and one part-time coordinator (£16.5K) and one data entry clerk (£16.5K) was calculated at £66K (£1.00 = approx. \$2.00). There were five training days for 10 project champions ($40 \text{ hrs} \times 10 \times £23.00 \text{ p.h} = £9,200$) and two coordinators ($40 \text{ hrs} \times 2 \times £17.00 \text{ p.h} = £1,360$), one-day training for 230 observers conducted over 10 days ($8 \text{ hrs} \times 230 \times £17.00 \text{ p.h} = £31,280$), and one Functional Analysis training day with 30 people ($8 \text{ hrs} \times 30 \times £17.00 \text{ p.h} = £4,080$). For these 16 training days, “off-task” costs were calculated at £45,920. Within each of the interventions, 41 of 55 observers conducted daily observations for 15-20 minutes at an estimated cost of £113,843 ($£17.00 \text{ p.h}/3 [60 \text{ mins} \div 20 \text{ mins}] \times 41 \text{ observers} \times 7 \text{ days} \times 70 \text{ weeks}$). Forty-one 30-minute weekly team briefings cost an estimated £181,917 ($41 \times 9 \text{ employees} \times £17.00 \text{ p.h}/2 [60 \text{ mins} \div 30 \text{ mins}] \times 82 \text{ weeks}$) while the (165) 30-minute goal-setting sessions across the three interventions were estimated to cost £12,622 ($165 \times 9 \text{ employees} \times £17.00 \text{ p.h}/2 [60 \text{ mins} \div 30 \text{ mins}]$). These estimates suggest total “off-task” costs of £420,302. The costs of remedial actions are unknown.

The company did not track or calculate the actual injury costs incurred during the project (a common occurrence). Therefore, the following reflects best estimates using figures available from a variety of external sources. Using 1998 figures, the average cost of a lost-time injury in the UK was £25,556.00. Thus, an estimated cost-savings of £357,784 was achieved for the 14 lost-time injuries avoided through this project (1998-1999). Additional £114,000 cost-savings obtained from the total reduction in recordable injuries ($n = 76$) were estimated at £1,500 per incident (Confederation of British Industry, 1990). Thus, total estimated direct cost-savings achieved £471,784 or more. Deducting ‘off-task’ labor costs of £420,302, suggests the project saved the company approximately £51K.

These figures, however, are likely to be serious under-estimates as they exclude indirect injury costs which are estimated by the British Health and Safety Executive (equivalent of OSHA) to be 10 times those

of recovered insured costs. Neither do they account for the 12% increase in productivity that tends to accompany a 50% improvement in safety performance (Stewart & Townsend, 1999), and thought to provide an internal rate of return in excess of 30%.

Exploratory Analyses

The main purpose of this study was to explore the effects Observation Frequency (OF) and Visible Ongoing Support (VOS) might have on the observed percent safe score. As there are no specific hypotheses about the direction of the relationships between the variables, two-tailed tests for all Pearson product-moment correlations (*r*) are reported.

Safety Performance and Visible Ongoing Support. Correlations between safety performance and VOS were statistically significant for all three interventions (see Table 3). The coefficient of determination (*r*²) for these variables indicated common variance of 52% (*r* = 0.72, *p* < .001) during the first intervention; 34% (*r* = .58, *p* < 0.01) during the second intervention and 22% (*r* = 0.47, *p* < 0.05) during the third intervention. Although the magnitude of the relationships is very large, it is difficult to ascertain the direction of effect. Results of other studies (Robertson et al., 1999), however, support the notion VOS exerts a substantial impact on safety performance, but this study indicated this effect might diminish over time.

Safety Performance and Observation Frequency. Table 5 reports small non-significant correlations between safety performance and ob-

TABLE 5. Individual Correlations for All Variables by Intervention

	SP/VOS	<i>r</i> ²	<i>p</i> < .05	SP/OFR	<i>r</i> ²	<i>p</i> < .05	OFR/VOS	<i>r</i> ²	<i>p</i> < .05
Int.1	0.72	.52	.001	0.37	.14	n.s	0.21	.04	n.s
Int.2	0.58	.34	.01	0.36	.13	n.s	0.36	.13	n.s
Int.3	0.47	.22	.05	0.26	.07	n.s	0.40	.16	.05

SP = Safety Performance

servation frequency for all three interventions. The coefficient of determination (r^2) between Safety Performance and OF across all three interventions indicated only 7-14% of the variance is common to both factors (first intervention $r = 0.37$, n.s.); second intervention $r = 0.36$, n.s.); third intervention $r = 0.26$, n.s.). This suggests observation frequency *per se* was not necessarily related to the safety performance of those employees observed. However, increased numbers of observations (i.e., daily) may help determine the reliability (Bausell, 1986) of the observed percent safe scores and affect an observer's behavior (Alvero & Austin, 2004).

Observation Frequency and Visible Ongoing Support. Small non-significant correlations between Observation Frequency and Visible Ongoing Support were obtained in the first and second interventions (first intervention: $r = 0.21$, n.s; second intervention: $r = 0.36$, n.s.). A significant correlation was obtained for the third intervention ($r = 0.40$; $p < 0.05$). The coefficient of determination (r^2) for VOS and OF across all three interventions indicates common variance of between 4-16%. This result suggests VOS is not related to OF *per se* (Contrary to Cook & McSween, 2002) although the magnitude of this relationship does appear to increase over time.

DISCUSSION

Ongoing managerial support behaviors, reported by employee observers in this study, appeared to be important *mediating* variables for behavioral safety interventions. The study results were consistent with a behavioral study in the British construction industry (Robertson et al., 1999), which used behaviorally anchored rating scales. However, the correlations obtained in the present study are of a larger magnitude which may have been due to the different measurement methods (i.e., perceptual ratings vs. behavioral measurement) used in combination with situational differences between a fluid (construction) and static (manufacturing) environment. The study results also supported the wider literature asserting

management commitment is essential for organizational interventions to succeed (Cohen, 1977; Rodgers et al., 1993).

The correlations indicated the association between VOS and the percent safe score diminished over time. It may be argued this is simply a statistical artifact related to possible range restrictions rather than a diminished relationship (i.e., the mean percentage scores for both variables are high in the third intervention, which may have lowered the correlation). This latter aspect appears unlikely as Pearson correlations take into account the *direction* for each set of scores, but are unrelated to the size of either mean score on which it is computed.

In this study, a perusal of the graphs shows percent safe scores increase, while VOS scores simultaneously decrease (toward the end of interventions two and three). Thus, each set of scores diverge. However, Pearson correlations are also based on standardized scores, which are used to test for the relative distance from their sample mean score (i.e., variance). All other things being equal, the greater the variability, the greater the size of correlation (Glass & Hopkins, 1996). Calculation of the coefficient of variation (the standard deviation divided by the mean, multiplied by 100) expresses variability as a percentage of the mean. This showed variance shrinkage for both percent safe (Intervention 1: 10.34%; Intervention 2: 10.23%; Intervention 3: 8.76%) and VOS (Intervention 1: 21.34%; Intervention 2: 8.67%; Intervention 3: 6.97%) over the three interventions, which may partially explain the reductions in size of correlation across time.

A partial explanation may also reside in managerial time constraints in combination with the percent safe score reaching levels in the mid-1990s. This may have led managers to question the need to continue high levels of support. The diminishing levels of project champion support over the three interventions may also have contributed to this.

The apparent diminishing effect of VOS on behavior might also involve the remedial action rate. VOS largely represents management supportive behavior in one form or another (i.e., accompanying observers, allowing time for feedback meetings, or fixing things). During the study, the site completed 151 of 216 actionable items (at a fairly constant rate across the three interventions). It seems reasonable to speculate,

therefore, that addressing these situational items (i.e., reducing unsafe conditions) positively influenced safety behavior by reducing the opportunity for people to behave unsafely.

Over a period of time, the fix rate may have reached a ceiling whereby system fixes would no longer exert the same impact on behavior that may have been experienced in the early stages of the project. At the same time, this reduced opportunity for fixing things may have been reflected in lower recorded weekly VOS levels, which in turn appears to demonstrate a diminishing effect of VOS on safety performance.

A complementary explanation for the overall strength of relationship between VOS and safety performance was formal weekly feedback. One VOS behavior monitored was whether weekly feedback briefings took place for each workgroup. One of the most robust findings in the behavioral safety literature is the effect of feedback on performance, whether graphic (Duff et al., 1993), verbal (Zohar, Cohen, & Azar, 1980) or written (Cooper et al., 1994). Because regular weekly briefings took place reliably about 90% of the time, the study findings could reflect the effects of the feedback delivered during the meetings (Algera, 1990) rather than management commitment *per se* (Robertson et al., 1999; Rodgers et al., 1993).

It would be useful if future research simultaneously assessed the effects of each feedback mode and individual components of management's commitment (e.g., remedial action rates, observer support) to help determine which variables exert the greatest influence on behavioral safety performance *per se*.

Commitment and Observation Relationships

Although there was a consistent relationship between Visible Ongoing Support (VOS) and observed safety performance, this was not the case between VOS and Observation Frequency (OF). Expectations were that VOS would help maintain the daily observation frequency levels requested (Cook & McSween, 2000). Instead, OF was substantially higher for the first intervention when there were lower levels of managerial commitment. Observation rates then reduced in the second intervention by

about 12% points and stabilized at this level for the third intervention at the same time as management commitment was increasing substantially. This convergence could provide a simple explanation why it appears the VOS/OF relationship strengthens as the VOS/Performance relationship weakens.

Unlike the Cook and McSween (2000) study which found that high supervisory involvement increased the percentage of employees conducting observations, this study found the opposite (i.e., despite VOS increasing, OF decreased during interventions two and three). The difference is likely related to study design: Supervisors were observers in their own right in Cook and McSween's study. In this study, a line-manager was asked only to conduct a paired observation with an existing employee observer once per week. Thus, there was a fundamental difference in the rationale and processes involved. However, it would be useful if future research extended the VOS measures to include many other types of managerial support behaviors (e.g., encouraging observations), and ascertain their effect on observation frequency.

Observation Frequency

Komaki et al. (1978) suggested three observations a week are sufficient to represent an accurate picture of safety behavior, but presented no empirical evidence as to why this is the case. Some data presented in various commercial behavioral safety literatures suggest there is a strong relationship with high levels of OF and injury reduction, but such data have yet to be reported in the scientific literature. This study found a very weak relationship (see the section below discussing injury reduction).

This study also failed to find any statistically significant relationships between OF and Percent Safe. In part, this may be due to methodological factors. The relationship is likely to be much stronger when there are many employees conducting "peer-to-peer" observations. However, the approach adopted here used individual observers who monitored their workgroups on a daily basis. This has probably contributed to insufficient variation in Observation Frequency between interventions.

Future research comparing the two approaches while systematically manipulating OF would be useful.

Nonetheless, some relationship was expected, given the increased opportunities presented for employees to experience individual consequences (e.g., verbal feedback) after a fellow employee's observation. Correcting or positively reinforcing safety behavior at the time of occurrence significantly improves performance (Fellner & Sulzer-Azaroff, 1984, Hopkins et al., 1986; Zohar et al., 1980). Unfortunately, measurement of the observer's frequency of verbal feedback did not occur on a weekly basis. However, self-report anecdotal evidence collected during a behavioral safety process audit by the author indicated approximately 50% of observers were giving regular verbal feedback to their colleagues at the point of observation. This audit finding may partially be reflected in the coefficients of determination (r^2) obtained in each intervention, suggesting there is a small non-significant overlap between Observation Frequency and percent safe: observation frequency helps determine the amount of verbal feedback provided. In turn, the verbal feedback helps to improve the levels of percent safe.

One explanation for OF differences across the interventions is that the initial observers were mostly union safety representatives who had been actively engaged in trying to improve safety prior to the study, and had previously demonstrated they were more committed to safety than others were. Observers from the later interventions were drawn from the workgroups, and had not previously been actively engaged in prior safety improvement activities. Hence, their underlying commitment to safety could have been somewhat lower. Further research focusing on observer commitment to a behavioral safety process should help to clarify the effects on both safety performance and Observation Frequency.

Injury Reductions

The decreases in injuries reported in this study are in line with those reported in Krause et al. (1999) for the first five years of operation, but less than that found in other studies (e.g., Cooper et al., 1994; Fellner & Sulzer-Azaroff, 1985; Laitinen et al., 1999). Explanations for this may

reside in the plant maintenance shutdowns that occurred during the study. For example, abrupt decreases occurring in the percent safe score during the first and last interventions were due to an extra contractor workforce arriving on site (for up to two weeks) from another European country (the machinery was built in this other country, and their personnel were the only ones trained and authorized to maintain it). Any injuries experienced by these contractors were recorded as though they occurred to directly employed personnel. Three percent of these contractors were involved in accidents during the shutdowns. Alternative explanations may reside in the different prevailing safety cultures (Cooper, 2000) of other study locations, although the impact of safety culture, or readiness assessments on behavioral safety processes is another issue still awaiting scientific enquiry (Cooper & Phillips, 2004).

In this study small, negative non-significant correlations were also obtained between the number of injuries, observation frequency, VOS, and percent safe. Cumulatively, all three variables only accounted for some 12% of the overall variance in the number of injuries. However, the percent safe result mirrors that obtained in Cooper et al. (1994), but is contrary to the findings of Reber and Wallin (1983) and received wisdom. Cooper et al. found much stronger relationships for sickness absenteeism and machine downtime, but did not examine observation frequency or VOS. In this study, the non-significant relationships between the number of injuries, and OF and VOS suggest that some other factor not yet measured, is responsible for actual injury reductions. It is recognized that the company's other safety initiatives would also have exerted some impact on injuries. Nonetheless, the lack of a clear inverse relationship between the percent safe scores and injuries is particularly puzzling: If approximately 90% of injuries involve employee behavior as is often asserted (e.g., Heinrich, 1931), then a strong relationship should be obtained when an intervention is focused on those injury-causing behaviors. Potentially, the remedial action rate and/or frequency of feedback and/or combinations of feedback mechanisms (e.g., graphical, verbal, and written) may exert strong mediating effects on the safe behavior and injury relationship, but the issue remains to be examined scientifically.

Limitations

Although an employee-led implementation is a significant strength of this study, a major scientific limitation relates to the lack of formal inter-observer agreement data. Attempts to calculate inter-observer reliability met with strong resistance from the company. For them, the important factor was employees were getting involved in safety *per se*. They did not want to create the impression that line-managers or others were overseeing the accuracy of observation, as this could reflect a lack of trust. However, VOS measures did evaluate the extent to which line-managers and the project team accompanied observers. This tactic was intended to help ensure observation accuracy as inconsistencies could be more easily highlighted and discussed during such paired observations.

The company also wanted to minimize the possibility of “faked” observations. Although the possibility of “faking” remains, almost all the observers would have to have consistently faked their observations over the whole study period to affect the study results. Because the timing of accompanied observations was unpredictable, (i.e., managers did it when they could find the time in their busy schedules) it is unlikely there was any systematic attempt at “faking.” Additionally, project coordinators conducted post-hoc comparisons on between-shift observation data in a deliberate attempt to look for potential faking.

Wherever possible, however, future employee-led studies should explicitly calculate inter-observer reliability scores to assess the extent to which they co-vary with the percent safe scores. It may also prove useful to conduct weekly line-management inspections to provide independent evidence of accuracy in the observed safety performance scores.

It could be argued the study did not allow a true return to baseline between interventions as the observation process itself is an independent variable. However, as indicated by Duff, Robertson, Cooper, and Phillips (1993) in a 12-month large-scale study on six construction sites utilizing a true ABABA multiple-baseline design (e.g., Komaki & Goltz, 2001) on each site for three categories of safety behavior introduced four weeks apart, performance does not necessarily return to pre-intervention levels

during withdrawal periods (Komaki et al., 1978). Rather, a time lag factor may lead to continuing performance increase for a few weeks before diminishing. In this study, therefore, a performance “dip” was expected after the four weeks “baseline” period, not during the baseline period itself.

It should also be noted that different observers and checklists were used for each of the interventions, and the behaviors constituting the dependent variable changed across the three interventions. Although a scientific limitation, these changes are desirable from a practical point of view. Employees indicate that four to six months is sufficient time for them to be an observer, and that someone else in the workgroup should become the workgroup observer as active observers “retire.” Equally, when continuously recording a particular behavior as 100% safe, there is not much point in keeping it on a checklist *ad infinitum*. In most workplaces, many other unsafe behaviors need addressing, particularly with the introduction of new work methods, materials technologies, and equipment. Thus, the changes in personnel and checklists are a deliberate strategy to help maintain the process over the longer term by ensuring periodic refreshment of the system. This may be an important aspect of employee-led behavioral-safety systems with a goal to improve people’s safety behavior in the workplace and reduce the company’s injury rates. This study appears to have succeeded in both these respects. However, there is an argument that these changes confound the OF and observed percent safe results, respectively.

Further limitations relate to Visible Ongoing Support. Although this was a first attempt, it can be argued the VOS measure used in this study was a weak proxy measure of management commitment and support. First, a simple yes/no response format presents problems associated with restriction of range in the scores. Second, some of the items may have been vague which allowed for different interpretations among observers. Third, VOS was only measured once per week. Greater variation in scores may have been achieved if this had been measured daily, or the number of instances of support during the week was recorded. Fourth, the lack of feedback to managers concerning VOS may have affected the levels of support demonstrated.

Conceivably, the effects of VOS on performance could have been greater and more consistent had feedback been provided. The project coordinator merely used the data to ensure events were occurring as planned and attempted to encourage those people who did not exhibit the requisite supportive behaviors. Future research examining VOS might usefully develop measures using a greater range of specific support behaviors that independently target senior, middle, and front-line management levels; extend the range of scores; measure support on a more frequent basis; ask managers to set VOS targets and provide regular VOS feedback.

CONCLUSION

The study results add to the literature by highlighting the apparent importance of management's commitment on behavioral safety performance. The study also provides some insight into the complexity of employee-led implementations that give rise to some serious issues awaiting scientific inquiry. There is a sufficiently robust body of evidence to show that behavioral safety is effective, but there is a distinct lack of evidence available about individual components of these interventions and the interactions between them. The lack of empirical evidence regarding the most fundamental component (i.e., observation) reveals the extent of this paucity of scientifically relevant applied research.

REFERENCES

- Algera, J.A. (1990). Feedback systems in organisations. In C.L. Cooper & I.T. Robertson (Eds.), *International review of industrial and organizational psychology, Vol 5*, (pp. 169-193). Chichester: John Wiley and Sons.
- Alvero, A.M., & Austin, J. (2004). The effects of conducting behavioral observations on the behavior of the observer. *Journal of Applied Behavior Analysis, 4*(4), 457-468.

- Austin, J. (2000). Some thoughts on the field of organizational behavior management. *Journal of Organizational Behavior Management*, 20(3/4), 191-202.
- Bandura, A. (1986). *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice Hall.
- Bausell, R.B. (1986). *A practical guide to conducting empirical research*. NY: Harper & Row.
- Buchan, R., McAra, D., Milne, A., & Heneghan, P. (1999) Drilling safety leadership: An experiment in experiential learning creates opportunities for enlightenment. *Offshore Europe Conference, Aberdeen, Scotland. 7-9 Sept. Society of Petroleum Engineers, SPE 56979*.
- Cohen, A. (1977). Factors in successful occupational safety programs. *Journal of Safety Research*, 9(4), 168-178.
- Confederation of British Industry (1990). *Developing a safety culture*. London: Confederation of British Industry.
- Cook, S., & McSween, T. (2000). The role of supervisors in behavioral safety observations. *Professional Safety*, 45(10), 33-36.
- Cooper, M.D. (1993). Goal-setting for safety. *The Safety and Health Practitioner*, 11, 32-37.
- Cooper, M.D. (1998). *Improving safety culture: A practical guide*. Chichester: John Wiley and Sons.
- Cooper, M.D. (2000). Towards a model of safety culture. *Safety Science*, 32(6), 111-136.
- Cooper, M.D. (2003). Behavior-based safety: Still a viable strategy. *Safety & Health*, National Safety Council, April, pp. 46-48.
- Cooper, M.D, Bissett, M.P., & Walters, S. (1999). *The B-Safe tracking program (2nd ed.)*. Applied Behavioural Sciences, Hull, UK.
- Cooper, M.D., & Phillips, R.A. (2004). Exploratory analysis of the safety climate and safety-behavior relationship, *Journal of Safety Research*, 35, 497-512.
- Cooper, M.D., Phillips, R.A., Sutherland, V.J., & Makin, P.J. (1994). Reducing accidents with goal-setting and feedback: A field study. *Journal of Occupational and Organizational Psychology*, 67, 219-240.
- Cooper, M.D., Robertson, I.T., Duff, A.R., & Phillips, R.A. (1992, January). *An examination of participative and assigned goal-setting in relation to the improvement of safety on construction sites*. Paper presented at the annual British Psychological Society's Occupational Psychology, Conference. Liverpool.
- Daniels, A. C., (1989). *'Performance Management: Improving Quality Productivity through Positive Reinforcement' (3rd ed)*. Georgia: Performance Management Publications.
- Duff, A.R., Robertson, I.T., Cooper, M.D., & Phillips, R.A. (1993). *Improving safety on construction sites by changing personnel behavior*. Health & Safety Executive Report Series CRR 51/93. London: Her Majesty's Stationary Office.

- Fellner, D.J., & Sulzer-Azaroff, B. (1984) Increasing industrial safety practices and conditions through posted feedback. *Journal of Safety Research*, 15, 7-21.
- Fellner, D.J., & Sulzer-Azaroff, B. (1985). Occupational safety: Assessing the impact of adding assigned or participative goal-setting. *Journal of Organizational Behavior Management*, 7, 3-23.
- Fox, D.K., Hopkins, B.L., & Anger, W.K. (1987). The long-term effects of a token economy on safety performance in open-pit mining. *Journal of Applied Behavior Analysis*, 20, 215-224.
- Frederick, J., & Lessin, N. (2000). Blame the worker: The rise of behavioral-based safety programs. *Multinational Monitor*, 21, 10-14.
- Glass, G.V., & Hopkins, K.D. (1996). *Statistical methods in education and psychology (3rd ed.)*. Boston: Allyn and Bacon.
- Grindle, A.C., Dickinson, A.M., & Boettcher, W. (2000). Behavioral safety research in manufacturing settings: A review of the literature. *Journal of Organizational Behavior Management*, 20(1), 29-68.
- Harper, A.C., Cordery, J.L., de Klerk, N.H., Sevastos, P., Geelhoed, E., Gunson, C., Robinson, L., Sutherland, M., Osborn, D., & Colquhoun, J. (1997). Curtin safety trial: Managerial behavior and program effectiveness. *Safety Science*, 24(3), 173-179.
- Heinrich, H.W. (1931) *Industrial accident prevention*. New York: McGraw-Hill.
- Hollenbeck, J.R., Klein, H.J., O'Leary, A.M., & Wright, P.M. (1989). Investigation of the construct validity of a self-report measure of goal commitment. *Journal of Applied Psychology*, 74, 951-956.
- Hopkins, B. L., Conrad, R. J., Dangel, R. F., Fitch, H.G., Smith, M.J., & Anger, W.K. (1986). Behavioral technology for reducing occupational exposures to styrene. *Journal of Applied Behavioral Analysis*, 19(1), 3-11.
- Kazdin, A. (1982). *Single-case research designs: Methods for clinical and applied settings*. New York: Oxford University Press.
- Komaki, J. L., Barwick, K.D., & Scott, L.R. (1978). A behavioral approach to occupational safety: Pinpointing and reinforcing safe performance in a food manufacturing plant. *Journal of Applied Psychology*, 63, 434-445.
- Komaki, J.L., & Goltz, S.M. (2001). Within-group research designs: Going beyond program evaluation questions. In: C.M. Johnson, W. K Redmon, T.C. Mawhinney (Eds). *Handbook of organizational performance: Behavior analysis and management*. (pp. 81-123). New York: Haworth Press.
- Komaki, J. L., Zlotnick, S., & Jensen, M. (1986). Development of an operant-based taxonomy and observational index of supervisory behavior. *Journal of Applied Psychology*, 71, 260-269.
- Krause, T. (1995). *Employee-driven systems for safe behavior*. NY: John Wiley and Sons.
- Krause, T. R., Seymour, K.J., & Sloat, K.C.M. (1999). Long-term evaluation of a behavior-based method for improving safety performance: A meta-analysis of 73 interrupted time-series replications. *Safety Science*, 32, 1-18.

- Laitinen, H., Marjamaki, M., & Paivarinta, K. (1999). The validity of the TR safety observation method on building construction, *Accident Analysis and Prevention*, 31, 463-472.
- Locke, E.A., & Latham, G.P. (1990). *A theory of goal-setting and task performance*. London: Prentice-Hall.
- Mattila, M., & Hyödynmaa, M. (1988). Promoting job safety in building: An experiment on the behavior analysis approach. *Journal of Occupational Injuries*, 9, 255-267.
- McAfee, R.B., & Winn, A.R. (1989). The use of incentives/feedback to enhance work place safety: A critique of the literature. *Journal of Safety Research*, 20, 7-19.
- Reber, R.A., Wallin, J.A., & Chhokar, J. (1990). Improving safety performance with goal-setting and feedback. *Human Performance*, 3, 51-61.
- Robertson, I.T., Duff, A.R., Marsh, T.W., Phillips, R.A., Weyman, A.K., & Cooper, M.D. (1999). *Improving safety on construction sites by changing personnel behavior: Phase 2*. Contract Research Report 229/1999. Sudbury: HSE Books.
- Rodgers, R., Hunter, J.E., & Rogers, D.L. (1993). Influence of top management commitment on management program success. *Journal of Applied Psychology*, 78, 151-155.
- Salancik, G.R. (1977). Commitment and the control of organizational behavior and belief. In B.M. Staw & G.R. Salancik (Eds.) *New directions in organizational behavior* (pp. 1-54). Chicago: St Clair Press.
- Stajkovic, A.D., & Luthans, F. (1997). A meta-analysis of the effects of organizational behavior modification on task performance, 1975-95. *Academy of Management Journal*, 40, 1122-1149.
- Stewart D.A., & Townsend, A.S. (1999). There is more to “health and safety is good business” than avoiding unplanned costs: A study into the link between safety performance and business performance. <http://www.behavioural-safety.com/articles>
- Zohar, D. (2000). A group-level model of safety climate: Testing the effect of group climate on micro-accidents in manufacturing jobs. *Journal of Applied Psychology*, 85(4), 487-596.
- Zohar, D., & Luria, G. (2003). The use of supervisory practices as leverage to improve safety behavior: A cross-level intervention model. *Journal of Safety Research*, 34(5), 567-577.
- Zohar, D., Cohen, A., & Azar, N. (1980). Promoting increased use of ear protectors in noise through information feedback. *Human Factors*, 22(1), 69-79.