Behavior-Based Feedback and Occupational Safety: Critical Impact of Social Comparison

Joshua H. Williams1 and E. Scott Geller

Virginia Polytechnic Institute and State University

Author Note: 1. Now at Safety Performance Solutions, Inc., 1007 N. Main St., Blacksburg, VA. 24060.
Abstract

The relative impact of global, specific, and social comparison feedback on safety behaviors was assessed at a large soft-drink bottling company. A 2 Feedback Type (Specific vs. Global) X 2 Social Comparison (Present, Not Present) analysis of covariance was used to test the hypothesis that specific, social comparison feedback would lead to the greatest improvement in percentage of safe behavior. Employees (n=97) received behavioral safety training and then developed a critical behavior checklist (CBC) for their work areas. Researchers used the CBC to a) systematically categorize work behaviors throughout the plant as safe or at-risk, and b) give employees written BB feedback each week. Social comparison feedback (SCF) led to significantly higher percent safe scores than did the no SCF conditions (mean percent safe was 78% vs. 68%, respectively). Follow-up Chi-Square analyses and practical considerations suggest global/SCF is optimal for improving safety performance. Limitations of the study and future implications for safety feedback research are discussed.
More than 11.3 million U.S. employees are seriously injured, and nearly 11,000 workers are killed on the job each year (Baker et al., 1992). It has been estimated that U.S. employers pay $155 billion in direct costs (e.g., workers' compensation, insurance premiums) associated with workplace injuries, amounting to over $1,400 per work-related injury (Baker et al., 1992; Miller, 1997; The National Committee for Injury Prevention and Control, 1989). Moving from at-risk to safe work habits is key to preventing injury (Califano, 1979; Geller, 1999).

Through the course of a workday, employees face numerous instances where they choose to work safely or at-risk. Unfortunately, workers are often rewarded for performing tasks in an at-risk manner because doing so is typically faster, easier, more comfortable, and more efficient or convenient than following the safe procedures (Geller, 1998a). Examples of at-risk choices include: not wearing protective gear, failing to follow standard energy-controlled lock-out procedures, lifting a heavy object without a hoist, standing on a machine instead of a ladder. In addition to such natural consequences like ease and comfort, some workers feel management pressure to take safety short-cuts for more efficient production.

The antecedent-behavior-consequence model of applied behavior analysis has been used frequently and successfully to prevent occupational injuries (e.g., Alavosius & Sulzer-Azaroff, 1986; Komaki, Collins, & Penn, 1978; Streff, Kalsher, & Geller, 1993). Antecedents or activators (e.g., safety signs) direct one’s focus and attention on relevant safety behaviors needed for a given task. Consequences (e.g., rewards) follow behavior and motivate the future occurrence of the rewarded behavior.
Behavioral approaches to injury control have a number of advantages, including:
a) they can be administered by people with minimal professional training (McSween, 1995); b) they can reach people in the natural setting where a problem occurs such as a particular work area (Geller, 1996); c) the leaders in these settings can be taught the behavior-change techniques most likely to work under specific circumstances (Baer, Wolf, & Risley, 1968, 1987; Daniels, 1989; Geller, 1998c); and d) they can be extremely cost-effective (Daniels, 1989; Sulzer-Azaroff & de Santamaria, 1980).

Improvements in safety-related behaviors following specific feedback have been demonstrated in a number of organizational settings including, for example: a plastics manufacturing plant (Sulzer-Azaroff & de Santamaria, 1980), a metal fabrications company (Zohar, Cohen & Azar, 1980), a bakery (Komaki, Barwick & Scott, 1978), a public work’s department (Komaki, Heinzmann, & Lawson, 1980), a university chemical laboratory (Sulzer-Azaroff, 1978), and a university cafeteria (Geller, Eason, Phillips, & Pierson, 1980). Fellner and Sulzer-Azaroff (1984) conclude that “Informational feedback on performance has been shown to be a simple, effective, and durable method for promoting safety” (p. 7).

In the safety research literature, behavioral feedback has been delivered in two different forms -- specific and global. Briefly, specific feedback is defined as the percentage of safe behavior occurrences over a given time period for certain target behaviors (e.g., bending knees when lifting is 45% safe for Week 1). Theoretical support for specific feedback is grounded in behavior analysis. Interventions to improve safety performance typically involve modifying or changing the salience of the antecedents and/or consequences of the specific safety behaviors targeted (Krause et al., 1996;
McSween, 1995). In fact, specific feedback has been demonstrated to be highly effective in various industry-based studies (e.g., Cohen & Jensen, 1984; Hopkins et al., 1986; Komaki et al., 1980; Ludwig & Geller, 1997).

In contrast, global behavioral feedback can be defined as an overall safety score based on the percentage of safe work practices over a given time period across a certain number of behaviors (e.g., 76% safe for Week 1). Support for using global feedback is driven by theories of generalized responding. Specifically, for a given set of behaviors, overall feedback without information regarding the specific behaviors contributing to the results should lead to more vigilant attention to all behaviors, resulting in generalized and long-term improvement (Boyce & Geller, 1999; Stokes & Baer, 1977). Global behavioral feedback has also been an effective intervention in numerous occupational safety studies (e.g., Austin et al., 1996; Cooper et al., 1994; Komaki et al., 1978; Reber, Wallin, & Chhokar, 1990). While there is ample evidence that both specific and global behavioral feedback lead to beneficial behavior change, no published research has systematically compared these two feedback approaches. As such, the present study compared the impact of giving employees specific versus global feedback regarding certain safety-related behaviors. In addition, the impact of social comparison feedback was evaluated.

Social comparison feedback (SCF) displays specific information about an individual’s or group’s performance in comparison to a relevant comparison group. Wood (1989) maintains, “…people not only wish to evaluate their abilities, they also feel pressure to continually improve them. When combined with the desire to compare with similar others, this drive upward leads the individual to strive toward a point slightly better than that of comparison others” (p. 232). Thus, SCF could serve a motivational
function. In other words, the addition of SCF, also referred to as normative feedback, should increase the impact of both specific and global behavioral feedback.

We could find no published research of SCF in the safety literature, however, there is empirical evidence that SCF improves performance with a word completion task (Mathieu & Button, 1992), a mental rotations task (Tindale, Kulik, & Scott, 1991), and a work-related simulation (Mitchell, Rothman, & Liden, 1985). However, some researchers suggest that SCF will have either no affect or negative impact on subsequent task performance because of group pressure to perform at the mean. Thus, exceptional performers may be motivated to perform less effectively. In fact, SCF led to diminished mean performance in a word search task (Johnson et al., 1996) and had no effect on actual job performance (Smither, Wohlers, & London, 1995). Thus, there is contradictory evidence regarding the influence of SCF on task performance and no published data on the influence of SCF on industrial safety performance.

In summary, there is a plethora of research suggesting behavior-focused feedback leads to increases in safe work practices and reductions in occupational injuries. However, less is known about the specific types of feedback which lead to the greatest improvement in safety performance. More specifically, the existing body of literature does not address the relative impact of specific versus global feedback and SCF versus no SCF on safety-related behavior. The current research made these comparisons.

Method

Participants and Setting

Participants were 97 employees from Shifts 1 and 2 of a soft-drink bottling facility located in southwestern Virginia. The majority of employees were male (69%)
and Caucasian (82%). Both cans and plastic bottles are filled with the product and
distributed via conveyers to large “labeling” machines, and then to large “packaging” (or
dematizers) machines. The facility consists of four production lines, a warehouse, and a
laboratory area. Because the daily activities of the laboratory are very dissimilar to the
other areas, laboratory observations were not used in the study. Although no fatalities
have occurred at this site within the last five years, 75 restricted workdays and 12 lost
days due to injury occurred in 1998. The majority of injuries are back strains/sprains and
hand/head abrasions.

Site-Specific Observation Checklist

The behaviors included on a critical behavior checklist (CBC) were chosen by the
safety manager and safety steering committee with guidance from the senior author.
Given the high percentage of back injuries, specific lifting behaviors were chosen.
Because of past near-misses and incidents involving fork-lift trucks, specific forktruck-
driving behaviors were also included on the CBC. In addition to protective gear, generic
safety concerns were chosen for the checklist, including stacking (danger of pallets
tipping over), handrail use (slips occur on stairs with wet surface areas from the
sanitation procedures), and overhead conveyer avoidance (causing head abrasions).

Site-Specific Areas

The employees were categorized into four similarly-sized groups of
approximately 25 employees each: 1) Group A = employees on Lines 1 and 2 of the
plant; 2) Group B = those on Line 3; 3) Group C = employees in the warehouse; and 4)
Group D = employees on Line 4. The tasks and requirements of the employees in these
four groups were relatively similar, although loading and lifting occur more frequently in
the warehouse. According to the safety director at the facility, injury reports were similar across groups.

Procedure

Behavioral safety education/training. The senior author delivered one hour of behavioral safety education/training to all employees, with approximately 25 participants per session. All training occurred on one day. Principles and practical applications of behavior-based safety were presented, with an emphasis on behavioral observation and feedback. Participants were informed that behavioral observations would follow training for ten weeks. This feedback was given individually to each employee, and participants were repeatedly asked not to discuss feedback information with coworkers. In other words, they were told to keep the feedback experiences personal and private. Also, participants were informed that observation data would not be shared with management, and that only group data would be presented.

Behavioral observations. After all employees received the behavioral safety education/training, behavioral observations were made for ten consecutive weeks. For Shifts 1 and 2, the same CBC was systematically used during two weeks of baseline, six weeks of intervention, and two weeks of follow-up. The CBC observations occurred every day, Monday through Friday. Observers who were unaware of the experimental conditions of the study followed a specific route for all sessions. When an employee was observed to be engaged in a task, the observers marked individual behaviors as either safe or at-risk, based on the operational definitions given on the back of the CBC. When a judgement about a specific behavior was not possible (as when long hair obscured a clear view of hearing protection), a line was drawn through the safe and at-risk box for that
specific observation to indicate a “non-observation.” The observations on both shifts lasted two to five minutes per target individual, with the length of an entire observation session ranging from 1 to 2 hours.

**Feedback intervention.** At the end of each week of the intervention period, feedback was presented graphically on bar graphs in sealed envelopes addressed to each individual employee. The specific target and non-target safety-related behaviors were listed on the feedback page above the graphs. Participants in the global feedback condition were given a single percent safe score for the week, whereas participants in the specific feedback condition received graphs for each target behavior.

Participants in Groups A (n=23) and B (n=22) received their area’s percent safe score(s) for the week compared to the group’s performance of the same area on the other shift. These percentages were global for Shift 1 and specific for Shift 2. Participants in Groups C (n=28) and D (n=24) only received their own group’s percent safe scores.

**Results**

**Reliability of Observations**

For 27% of the 15,183 total behavioral observations and balanced across phases, two trained and experienced research assistants independently observed the participants. These observers communicated with each other during the observation procedures only to identify the person being observed. After identifying a target individual, each observer completed the CBC independently. A percent agreement score was calculated for both safe and at-risk behaviors for each behavioral category with the formula: Total Number of Agreements/Total Number of Agreements + Total Disagreements X 100. The mean agreement percentages for safe observations were 93% (n=1725) for PPE, 80% (n=3564)
for Lifting, 77% (n=3880) for Forktruck Driving, 91% (n=975) for General Safety Behaviors. The mean agreement percentages for at-risk observations were 88% (n=681) for PPE, 83% (n=1502) for Lifting, 78% (n=2607) for Forktruck Driving, and 93% (n=313) for General Safety Behaviors (cutting, conveyor avoidance, stacking).

**Percent Safe Scores**

Daily data from CBC cards were converted to daily percent safe scores (% Safe Score = Total Safe Observations/Total Safe Observations + At-Risk Observations X 100) across all behaviors for a single daily mean percentage (i.e., dependent variable). These means were collapsed over a three-day period for the structured analysis because there were large variations in the number of observations made per individual each day. This daily number ranged from 1 to 223, with a mean of 78 and standard deviation of 49.7. This three-day measurement unit was used for the 2 Feedback Type (Global or Specific) X 2 Social Comparison (Present or Not Present) analysis of covariance (ANCOVA), using baseline scores as the covariate.

The ANCOVA indicated that safety feedback was effective at increasing percent safe scores over baseline. Overall, mean percent safe scores across all feedback conditions rose from 64% at baseline to 73% at intervention (and fell only slightly to 72% at follow-up). A plot of mean percent safe scores per feedback conditions at baseline, intervention, and follow-up is provided in Figure 1.

------------------------

Insert Figure 1 about here

------------------------
Global versus Specific Feedback

Mean percent safe scores at intervention were only slightly higher in the specific (M = .74, SD = 8.7) versus global (M = .72, SD = 9.7) feedback conditions. Specific/SCF led to higher percent safe scores than Global/SCF (M = .79 and .77, respectively) and Specific/NoSCF led to higher percent safe scores than Global/noSCF (M = .70 and .66, respectively). Overall, the main effect for feedback level was not statistically significant (F = .17, p > .05).

Social Comparison versus No Social Comparison Feedback

Percent safe scores were substantially higher in the SCF (M = .78, SD = 6.94) versus no SCF conditions (M = .68, SD = 8.43). Global/SCF led to a higher mean percent safe score than Global/noSCF (M = .77 and .66, respectively). Similarly, Specific/SCF led to a higher mean percent safe score than Specific/noSCF (M = .79 and .70, respectively). Not surprisingly, the main effect for feedback type was statistically significant (F = 10.8, p < .05).

Feedback Level by Feedback Type

Specific/SCF resulted in the highest percent safe score across feedback conditions (M = .79, SD = 6.7), followed by Global/SCF (M = .77, SD = 7.41), Specific/NoSCF (M = .70, SD = 8.48) and Global/NoSCF (M = .66, SD = 8.24). Overall, the Feedback Level by Feedback Type interaction was not significant (F = .02, p > .05). A plot of the feedback level by type interaction is provided in Figure 1.

A series of Chi-Square analyses were used to test for significant differences between the four feedback conditions at intervention. Specific/SCF (M = .79, SD = 6.70) was not significantly higher ($X^2_{(1)} = .46, p > .05$) than Global/SCF (M = .77, SD = 7.41). Next,
Global/SCF ($M=.77$, $SD=7.41$) was significantly ($X^2_{(1)}=14.3$, $p<.05$) higher than Specific/NoSCF ($M=.70$, $SD=8.48$). Global/NoSCF ($M=.66$, $SD=8.24$) was significantly lower ($X^2_{(1)}=12.81$, $p<.05$) than Specific/NoSCF ($M=.70$, $SD=8.48$). With follow-up data, only the percent safe score for the Global/NoSCF dropped significantly ($X^2_{(1)}=8.54$, $p<.05$) from intervention to follow-up.

Discussion

Performance feedback research seldom includes experimental comparisons of the specific effects of the different feedback characteristics (Balcazar, Hopkins, & Suarez, 1985/86). This lack of research extends to the safety literature, which consistently demonstrates that behavior-based feedback improves safety performance without exploring the characteristics of feedback driving these results. To this end, the objective of the current study was to add to the existing literature by increasing our understanding of the relative impact of global, specific, and social comparison feedback on subsequent safety performance.

Similar to prior industrial safety research using behavior-based (BB) feedback as an intervention (Fellner & Sulzer-Azaroff, 1984; Sulzer-Azaroff & de Santamaria, 1980; Zohar, Cohen & Azar, 1980), behavioral feedback in the current study led to clear increases in safe work practices. Based on more than 15,000 observations over an eight-week period, overall percent safe scores increased from 64% to 73% following the introduction of BB feedback and was maintained during the withdrawal period (72%).

The results supported the use of social comparison feedback to improve safety performance. Prominent differences in percent safe scores between the SCF and NoSCF conditions at intervention are reflected in Figure 1. Overall, SCF led to an overall 10
percentage point increase in percent safe scores over the NoSCF conditions. More specifically, the intervention impact was significantly greater for the Specific/SCF condition than Specific/NoSCF, and the feedback benefits were significantly higher for Global/SCF than Global/NoSCF.

Although social comparison feedback was beneficial at improving safety performance, the expected advantage of specific over global BB feedback was not found. Specifically, the impact of Specific/SCF and Global/SCF was similar. On the other hand, Specific/NoSCF led to a significantly higher mean percent safe score than Global/NoSCF, and only the Global/NoSCF group showed a decrease in percentage of safe behaviors after the feedback intervention was withdrawn. This suggests specific feedback is superior to global feedback when SCF is not present.

While specific feedback may outperform global feedback in the absence of SCF, the similar impact of Global/SCF and Specific/SCF feedback has practical implications. It took an extremely long time to encode, calculate, aggregate (by week) and the post percent safe scores for the nine specific behaviors targeted in this research. This was especially true when SCF was added, and specific feedback was given to more than one group. Specifically, it took the senior author five to six hours a week to prepare the charts for the specific feedback conditions. Conversely, computing one single overall score (i.e., global feedback) even with SCF, took only about 30 minutes a week to complete. The difference in response cost between the two feedback conditions week after week cannot be overstated.
Most safety professionals have limited time and resources for implementing and maintaining the variety of interventions available to improve safety performance. In an effort to reduce injuries, many professionals choose BB feedback to supplement existing safety programs. The results of the current study suggest that Global/SCF is basically as effective as Specific/SCF in improving safety performance, at a fraction of the time and effort required to provide it. For safety professionals without SCF) led to no meaningful change over baseline and a drop in safe work practices after removal of BB feedback. Simply put, global feedback by itself had almost no effect on subsequent safety performance, and given the follow-up results, may have actually done more harm than good. In contrast, substantial improvements from baseline to intervention were shown with the other three feedback conditions.

While providing employees one global score did not improve performance, adding a second overall score (from another shift) led to a feedback effect comparable to the two specific feedback conditions. Apparently, providing comparison information affected the workers’ collective desire to outperform similar other, and thereby served a motivational function.

The powerful influence of SCF is not surprising given past theoretical and empirical support, including Festinger’s (1955) seminal article, “A Theory of Social Comparison Processes.” Festinger (1955) argues that in Western cultures individuals are driven to continually improve performance and seek out similar other with whom to gauge their performance. Thus, the critical factor in the evaluation of an individual’s performance is his/her relative standing in comparison to others.
Further, Festinger’s (1955) concept of the unidirectional drive upward suggests that individuals are motivated to not only improve their own performance, but also to outperform the comparison group (Wood, 1989). The result of this comparison information is improved performance, as demonstrated by a number of empirical studies (e.g., Goltz et al., 1989; Mathieu & Button, 1992; Mitchell, Rothman, & Liden, 1985; Tindale, Kulik, & Scott, 1991).

In the current study, SCF likely: a) served as a motivational referent for comparison; and b) triggered participants’ collective desire to outperform “similar others”. The result is that participants in the SCF demonstrated significant and substantial improvements over baseline.

Geller (1999) proposed that behavior-change interventions are instructional, motivational, or supportive, and the impact of a particular intervention depends upon the awareness versus motivational state of the target individuals. Since, Global/SCF was just as effective as Specific/SCF, it can be assumed the employees did not need specific direction or instruction (as provided by specific feedback). Rather, the employees apparently knew how to perform their jobs safely but needed some extrinsic motivation to follow the non-safe policies implied by the nine target behaviors. Equivalent motivation was provided by a global comparisons score and by specific feedback reflecting an accountability system for nine different behaviors.

Conclusion

Behavior-based feedback increased percent safe scores from baseline to intervention, and these improvements were maintained for three of four groups during follow-up. More specifically, Global/SCF provided the most promising results in terms of
how to optimize behavioral feedback. Conversely, Global/NoSCF resulted in the least promising results for behavioral feedback presentation in organizational settings.

Unfortunately, the most common type of behavioral feedback in applied settings is a single, plant-wide percent safe score (Krause, Hidley, & Hodson, 1996). So, the majority of organizations using behavioral feedback are probably not experiencing the full potential and benefits of this intervention.
REFERENCES


Figure 1. Feedback Plot

- Baseline
- Intervention
- Follow-Up

- Global/NoSCF
- Global/SCF
- Specific/NoSCF
- Specific/SCF