

An Implementation of Protocol Analysis and the Silent Dog Method in the Area of Behavioral Safety

Alicia M. Alvero, Queens College at the City University of New York
John Austin, Western Michigan University

Recent research has demonstrated that conducting safety observations increases the safety performance of the observer. The purpose of this study was to help determine whether observers make self-verbalizations regarding their own safety performance and whether these reports are functionally related to safety performance. In order to answer these questions two experiments were conducted using both protocol analysis and the silent dog method. The objective of Experiment 1 was (a) to determine whether safety performance with continuous, concurrent talk-aloud procedures is functionally equivalent to safety performance without talk-aloud reports, and (b) to determine whether that safety performance is altered when participants are presented with a distracter task. The goal of Experiment 2 was to determine whether the safety-related verbalizations made by Experiment 1 participants were task-relevant and functionally related to safety performance. The results from both Experiments 1 and 2 provide support for the existence of a functional relationship between safety-related verbalizations and increases in safety performance.

Key words: behavioral safety, silent dog method, protocol analysis, behavioral observations, self-generated rules.

Behavior-based safety (BBS) is a proactive approach to improving safety within organizations that utilizes behavior analysis principles. The distinguishing feature between BBS processes and more traditional safety management approaches is a focus on critical behaviors that prevent injuries rather than focusing on accident rates and workplace conditions alone. Therefore, BBS is considered a *proactive* approach versus more traditional *reactive* approaches toward safety. The BBS approach aims to decrease the number of at-risk behaviors and increase the number of safe behaviors within an organization in order to decrease injuries (McSween, 2003).

Extensive research has identified the principal components of effective BBS processes

(Komaki, 1986; Komaki, Heinzmann, & Lawson, 1980; Sulzer-Azaroff & Fellner, 1984). Successful BBS processes normally include: assessment and identification of performance targets, development and implementation of a behavioral observation process, review of observation data, and implementation of a behavioral feedback process. The observation process involves training employees to conduct safety observations using a behavioral checklist. When conducting observations, observers (i.e., trained employees) approach other employees, observe, and score their performance using the behavioral checklist. Some research suggests that the observation process itself may serve as an effective tool in increasing the safety performance of the observer (Alvero & Austin, 2003, 2004). These findings are especially important because consultants have called for employee or research-driven safety programs (Krause, 1997; McSween, 2003), representing a shift from the earlier management-driven programs (for an example of these, see Komaki, Barwick, & Scott, 1978).

This shift to employee-driven programs raises questions about which activities are the most beneficial for employees to participate in. Some have argued that participating as an observer is critical for changing behavior (McSween, 2003). There are two fundamental questions related to conducting BBS observations. First, do observers perform more safely

The research presented in this manuscript was part of the first author's dissertation work at Western Michigan University, which was funded, in part, by the Western Michigan University Graduate College Travel and Research Grant. Some of the data from this paper were previously presented at the 29th Annual Convention of the Association for Behavior Analysis in San Francisco, CA.

The authors would like to extend many thanks to all of the research assistants who helped with the extensive data collection process. Alicia M. Alvero is now affiliated with the Department of Psychology at Queens College, and correspondence concerning this article should be addressed to her at Queens College, CUNY, Department of Psychology, 65-30 Kissena Blvd., Flushing, NY 11367; phone: 718-997-3212; email: alicia_alvero@qc.edu.

as a result of conducting observations? Second, if conducting observations changes behavior, how can we explain this effect? Recent research (Alvero & Austin, 2003, 2004; Sasson, 2002) has targeted the first question and demonstrated the existence of an observer effect: Conducting safety observations improves the safety performance of safety observers. The present study was an attempt to answer the second question; however, there were complex methodological and measurement issues involved in determining why observing the behavior of others affects the observer. Based on results of previous research (Alvero & Austin, 2004; Austin & Delaney, 1998; Mawhinney & Austin, 1999), it was speculated that understanding this issue would require understanding the various “thoughts” and “cognitions” of the observer. In other words, it would require examination of the sometimes covert verbalizations evoked before and after participants conduct observations of the behavior of others. To overcome this methodological challenge, protocol analysis and the silent dog method were used. These procedures are designed to collect data on typically unobservable phenomena.

Protocol analysis is a reliable and valid tool used in cognitive psychology to analyze the thoughts of a person as they perform a task (Austin & Delaney, 1998; Ericsson & Simon, 1993). The term “protocol analysis” has been used to describe a variety of methods for obtaining verbal reports. Throughout this manuscript, the term is used to refer to the set of procedures described by Ericsson and Simon, that is, the use of concurrent verbal reports. In this method, participants are trained to “think” or talk aloud while performing a task so that their concurrent, task-related “thoughts” or verbalizations may be recorded and later analyzed (for a detailed description of the methodology, please see Ericsson & Simon).

Although protocol analysis is a useful tool for obtaining the verbalizations that occur during task performance, two problems with the methods described by Ericsson and Simon (1993) apply to the present study. The first problem is related to the novelty of this research and the second to the variable of interest—specifically, the potential functional relationship between verbalizations and safety performance. Protocol analysis is most often used for problem-solving tasks or tasks that are well-defined

and have only one correct answer (Austin & Delaney, 1998; Ericsson & Simon). Therefore, verbalizations for each problem often fall into a pattern of common actions that frequently occur during problem-solving.

It is very unlikely that this will be the case for the present study for two reasons. First, the task of interest, safety performance, is not well-defined in terms of having only one correct answer. Undoubtedly, there is only one way to perform a task safely, but we are not likely to observe only one type of verbalization, or pattern of verbalizations, related to increases in safety performance. The second reason it is unlikely all protocol segments could be encoded as one of a small list of acts is because participants performed several tasks simultaneously (two of these tasks were measured: an assembly task while performing safely). Using protocol analysis to obtain important verbal information about this second “subtask” (safety performance) is novel. Therefore, it is not clear whether the above-described protocol analysis method is best designed to approach this novel experiment.

The second concern with using this traditionally cognitive approach is related to the research variable of interest. Protocol analysis is used to determine verbalizations that covary with correct performance, but the question of relevance is why conducting safety observations results in increases in safety performance. It was hypothesized that participants’ safety-related verbalizations are functionally related to changes in safety performance. Because of the differences between the typical applications of protocol analysis and the present study, a more behavioral approach to protocol analysis seemed more appropriate. Although a behavior analytic approach would not solve the concerns related to the novelty of this research, it would address the issue related to the functional relationship of interest.

A slightly modified version of protocol analysis, the “silent dog” method¹ of analyzing the impact of self-generated rules, allows

¹ The method is labeled the silent dog method, after the Sherlock Holmes’ mystery *Silver Blaze* in which Holmes knew the identity of a murderer because the dog in the stable had not barked when a horse was removed (Doyle, 1892/1986). In the silent dog method, it is the lack of a behavioral effect of self-report that suggests that performance is governed by a self-generated rule and that the verbal report is functionally similar to that rule (Hayes et al., 1998, p. 60).

researchers to determine the behavioral function of the thoughts and verbalizations that occur during task completion (Hayes, White, & Bissett, 1998). For this reason, the silent dog method can be used to help determine what “thoughts” or verbalizations (if any) are controlling participant safety behaviors after participants conduct safety observations on the behavior(s) of others. In other words, the silent dog method allows one to determine if conducting safety observations increases the observer’s safety performance as a result of the self-generation of rules, in part, or if the behavior change is simply contingency shaped (i.e., under the direct control of reinforcement and physiological cues).

There are three controls specified in the silent dog method and all three must be present to determine if verbal reports are functionally equivalent to rules, and thus, result in safety performance changes. First, it must be shown that performance on a task with continuous, concurrent talk-aloud procedures is functionally equivalent to performance without talk-aloud reports (i.e., it must be demonstrated that talking aloud during a task does not alter task performance). Secondly, it must be demonstrated that task performance is altered when participants are presented with a distracter task (i.e., performance is altered because the distracter prevents self-generation of rules or self-evaluative thoughts or for some other reason). The third control ensures that the absence of an effect between performance in the talk-aloud and silent condition (the result of control 1) is not because the task is contingency-shaped or that verbalizations are automatic or task irrelevant (Hayes et al., 1998). This is achieved by employing relevant elements of the verbal report (produced in control 1) as an external rule to new participants (for a detailed description of each control, please refer to Hayes et al.).

In summary, the purpose of this study was to help determine whether observers make self-verbalizations regarding their own safety performance and whether these reports are functionally related to safety performance. Protocol analysis was utilized to analyze the thoughts of a person as he or she performs a task and the “silent dog” method was used to help determine the behavioral function of the thoughts and verbalizations that occur during task performance.

EXPERIMENT 1

METHOD

The direct purpose of this manuscript is to present the methods and discuss an analysis of the verbal data. Therefore, the details of the experimental procedures used to manipulate safety performance (specifically, conducting safety observations) will not be discussed. For details regarding this information, (e.g., baseline procedures, intervention procedures, individual safety performance, etc.) contact the first author.

The purpose of Experiment 1 was to satisfy the first and second control requirements of the silent dog method, and thus, was twofold: (a) to show that safety performance with continuous, concurrent talk-aloud procedures is functionally equivalent to safety performance without talk-aloud reports, and (b) to demonstrate that safety performance is altered when participants were presented with a distracter task. The purpose of this second control was to further strengthen the hypothesis that safety-related verbalizations may be functionally related to increases in safety performance. In other words, if we suspect increases in safety performance are functionally related to safety-related verbalizations then we would expect a distracter task, which prevents the occurrence of safety-related verbalizations, to alter, or decrease, safety performance.

Participants, Setting and Sessions

Participants were eleven undergraduate students, between 18 and 35 years of age, at a midwestern university. All participants were compensated \$5.00 an hour for their participation throughout the study and were paid after each session. The study took place in a research lab, consisting of two observation rooms, located on the university campus. Each room was equipped with a video camera, table, chair, and materials required for a simple assembly task. All sessions lasted 15 minutes and were recorded so data could be analyzed at a later time.

Definition of Dependent Variables

Safety performance. The safety related target behaviors were all related to sitting pos-

ture. Government ergonomic reports were reviewed in order to determine the appropriate definition for proper sitting position (National Institute for Occupational Safety and Health, 1998; Office of Health and Safety Information System, 1998), and the definitions listed below were developed from these documents.

Sitting—(a) back upright, parallel to the back of the chair (not leaning at an angle against it), (b) shoulders in line with the back, not slouched forward, (c) both feet should be flat on the floor (ball of foot and heel should touch floor).

Safety-related verbalizations. A safety-related verbalization was defined as any vocal-verbalization that was related to safety and/or any of the target behaviors (feet, back, shoulder positions). For example, if a participant stated, “let me keep my feet flat,” “I should sit safely,” or “oops, I forgot to straighten my back,” it was counted as one safety-related verbalization. It is important to mention that these data were calculated only for the intervention sessions because the interest of this research is in the relationship between safety-related verbalizations and increases in safety that occur as a result of conducting observations (intervention). Therefore, safety-related verbalizations were not analyzed during the baseline and information phases because participants did not conduct observations during these phases.

Procedures

Talk-aloud and silent group. Before the start of the study, participants were randomly assigned to one of two groups: a talk-aloud or a silent group. Those in the talk-aloud group were trained to talk aloud by completing a practice session and were instructed to talk aloud during all phases of the study. Participants assigned to the silent group participated throughout the study without talking aloud during sessions.

Participants in the talk-aloud group were trained on how to talk or “think” aloud using a slightly modified version of the standard think-aloud instructions developed by Ericsson and Simon (1993). The standard procedures developed by Ericsson and Simon require participants to practice talking aloud while solving a set of problems. The nature of the task used in this experiment required participants to follow a set of instructions (place beads on a thread in

a specific color sequence), therefore, the talk-aloud training procedures involved presenting participants with various sets of instructions (e.g., placing papers in a specific order, sorting objects into containers, etc.) and asking them to practice talking aloud while following the set of instructions. The researcher prompted participants to continue talking aloud if they were silent for more than 5 seconds during the training. Participants practiced this procedure until they reported feeling comfortable talking aloud throughout the session and did not require any prompts by the researcher. A two-way beeping device was placed in the observation room during the session. The student investigator “beeped” the device if participants were silent for more than 5 seconds to remind them to continue talking aloud.

Baseline. As previously mentioned, all of the details of the experimental procedures will not be discussed, but a broad description of the baseline and intervention phase procedures may help with the general interpretation of the safety performance results. At the start of each baseline session, all participants were given instructions on how to perform an assembly task. The assembly task involved stringing beads onto a plastic string in a specific color order. This task was repeated throughout the duration of the session, thus trying to simulate the work a person might perform on an assembly line.

Observation. At the start of each observation phase session participants observed and scored someone else’s safety performance using a safety checklist which listed and defined each of the target behaviors. After conducting a safety observation, participants performed the assembly task described in baseline.

Observation plus distracter phase. After safety performance stabilized, participants in both the talk-aloud and silent conditions continued to conduct safety observations before each session, but they were also presented with a “cognitive distracter” (e.g., a “repeat after me” combined with a problem solving task). Participants were exposed to the distracter task after they conducted a safety observation and while they performed the assembly task. The distracter that was presented during this phase required participants to wear a headset attached to a handheld tape player, repeat what was said on the tape, and attempt to solve the problems presented on the tape (e.g., when presented with

the question, “What is 2 plus 5 plus 4 divided by 3?,” participants were required to repeat “What is 2 plus 5 plus 4 divided by 3” and provide an answer to the question). The purpose of the distracter was to prevent self-evaluation regarding safety performance, and as a result, decrease safety performance; thus, satisfying the second control requirement of the silent dog method.

All participants were instructed to talk-aloud during the distracter task phase to ensure that all participants were exposed to the distracter task. The researcher reviewed the videotaped distracter phase sessions to ensure that all participants engaged in the distracter task.

Safety Performance

Safety performance was measured using a 30-second momentary time-sampling procedure. A safety percentage was calculated by dividing the total number of intervals scored as “safe” by the total number of intervals scored as both “safe” and “unsafe” and multiplying by 100%.

Productivity Performance

Productivity performance was measured as a secondary variable. Productivity performance was defined as the number of beads thread onto the string in the correct color sequence during the 15-minute work session.

Verbal Data Analysis

Throughout the course of the experiment, verbalizations for all participants in the talk-aloud condition were transcribed into text. The text files were then divided into segments. A segment may be a sentence, clause, phrase or even a single word depending on the organization of a person’s oral prose (Ericsson & Simon, 1993). Segments were then randomly reordered, and presented out of context to coders who rated each segment into one of the following categories: (1) off-task statements (e.g., I have to go class, etc.), (2) environmental factors (e.g., this desk is too high; the room is too cold, etc.), (3) assembly task-related (e.g., orange, blue, green, I need to untangle this string, etc.), (4) target behavior-related (sitting like this

is so uncomfortable, my shoulders ache, etc.), (5) general safety-related statement (e.g., I should sit properly so I don’t ache, my feet are nailed to the floor, etc.), (6) specific safety-related statement (e.g., repetition of any part of the safety definitions provided on the safety checklists), and (7) general statements concerning the study (e.g., I wonder what this study is about, How much time do I have left, etc.). The coded segments were then reassembled in their original order for analysis.

The percentage of segments which were coded into each of the safety-related categories (e.g., general safety-related and specific-safety related) was calculated. After an analysis of the protocols, descriptions of safety rules were generated that were representative of the safety-related verbalizations that occurred concurrent with any increases observed in safety behavior for Control 1 talk-aloud participants.

Interobserver Agreement

Interobserver agreement (IOA) was conducted on all safety performance averages. As a reliability check, a second experimenter scored 30% of all sessions. Interobserver agreement between the two experimenters was calculated as follows: the number of agreements divided by the number of agreements plus disagreements multiplied by 100%. An agreement was defined as any occurrence in which both experimenters scored the same mark (safe or unsafe) for a behavior for each corresponding interval. IOA averaged 96.8% on participant safety performance.

IOA was also conducted on all transcriptions of talk-aloud participant sessions. At the end of the study, when all transcriptions were completed, an experimenter compared the written text with the audio statements and all discrepancies, or disagreements were noted. IOA was calculated by dividing the total number of agreements by the total number of agreements plus disagreements and multiplying by 100%. IOA averaged 99.5%. All discrepancies were reviewed by the first author and corrected. IOA was also calculated on all coded segments. Two experimenters coded all segments and discrepancies between data sets were noted. IOA averaged 99.85%. All discrepancies were reviewed by the first author and coded into the category that most closely described the contents of the segment.

Table 1
Participant safety performance across phases.

Target Behavior	Phase	Talk Aloud Group Participants						Silent Group Participants					
		0	1	2	3	5	6	7	8	9	10	11	
		Average Percent Safe						Average Percent Safe					
Back Position	Obs.	85.3	85.1	98.8	80.8	85.2	54.1	98.9	82.5	89.75	99.2	98.1	
	Distractor	15	8	40	12.5	9	25	70	46.5	38.5	36.5	15.7	
	Reversal to Obs.	93	100	95	*	97	67	90	93.5	93	100	100	
Shoulder Position	Obs.	81	82.3	95.7	76.6	76	48.6	98.9	78.5	89.75	99.5	98.1	
	Distractor	3.25	5.3	21	4.75	6.75	5	63.5	46.5	38.5	20	15.7	
	Reversal to Obs.	93	100	81.5	*	97	37	87	91.5	93	100	100	
Feet Position	Obs.	97.8	95	93.7	2.4	98	100	59.4	92	95.7	99.8	100	
	Distractor	50.75	57	73.7	3.25	20.75	71.5	0	16.5	96	95	95.7	
	Reversal to Obs.	97	100	96.5	*	100	100	100	98.5	100	100	100	

*Participant dropped out after the distractor phase.

RESULTS

Group Safety Performance Comparisons

Figure 1 shows average safety performance for both the talk-aloud and silent groups across the baseline and observation phases. Participants in the talk-aloud condition averaged 82% safe performance on the target behaviors during the observation phase, and participants in the silent condition averaged 89.5% safe performance. Visual inspection of the data do not seem to indicate any substantial differences exist between group safety performance.

Observation Plus Distracter Phase Performance

Table 1 summarizes average safety performance for each participant across three phases: observation phase, observation plus distracter phase, and a reversal to observation phase. Safety performance on all target behaviors decreased an average of 60% for the talk-aloud group, and 44% for the silent group with the presentation of the distracter task, and increased an average of 70% and 48% for each group, respectively, upon its removal.

The number of beads thread in the correct color sequence averaged 292 across all participants during the observation phase, and decreased to an average of 156 during the observation plus distracter phase. Productivity performance increased to an average of 318 during the reversal to observation phase.

Verbal Data

The number of segments which occurred during the observation phase and were coded into the safety-related categories totaled 132. The percentage of segments coded as general-safety related was 89.4, and 10.6% were coded as specific safety-related. During the distracter phase, 100% of verbalizations were related to the distracter task.

DISCUSSION

The goals of Experiment 1 were to satisfy the first and second control requirements of the silent dog method, and establish a framework for the methods, specifically for the independent variable, used in Experiment 2 (the goal of which was to satisfy the third control requirement). According to Hayes et al. (1998)

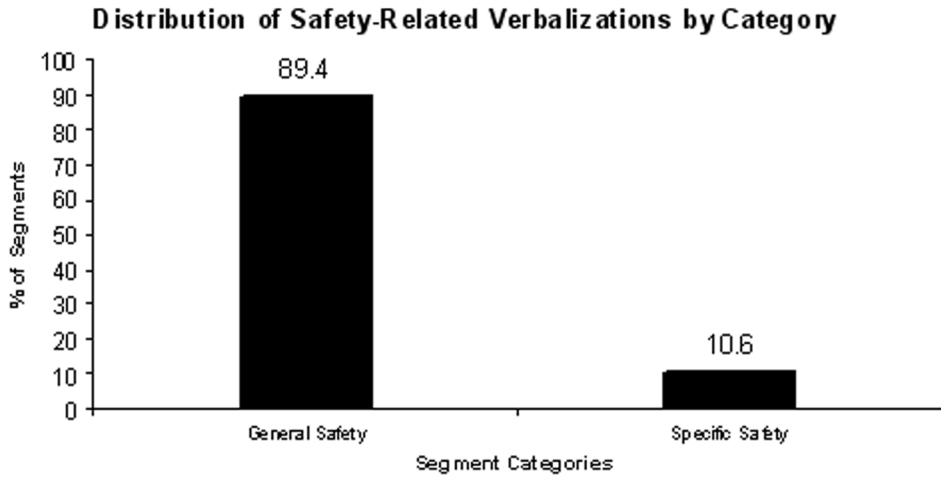


Fig. 1. Talk-aloud and silent group average safety performance.

all three controls must be present in order to determine the functional nature of self-rules or self-verbalizations, should any exist. The results from Experiment 1 suggest that observers make safety-related verbalizations as a result of conducting observations. The results also meet the second silent dog method control requirements, but whether the first control was met is unclear. According to the analysis posed by Hayes et al., analyzing the results within the context of the silent dog method requirements will help determine if the results of Experiment 1 suggest the existence of a functional relationship between safety-related verbalizations and increases in observer safety performance.

Silent Dog Method Control 1

The first silent dog method control states it must be shown that performance on a task with continuous, concurrent talk-aloud procedures is functionally equivalent to performance without talk-aloud reports (i.e., it must be demonstrated that talking aloud during a task does not alter performance). How one should demonstrate the equivalence between group performances is not specified. In other words, "How do we know that performance for talk-aloud participants is equivalent to the safety performance of silent group participants?" This is not a question that can easily be answered. An attempt was made toward finding an answer by visually analyzing safety performance data across both groups. Due to the experimental question of interest, group comparisons were

made only for safety performance data measured during the observation phase. The overall purpose of this research was to help determine the behavioral function of conducting safety observations. Therefore, the main variable of interest was the safety performance that resulted from conducting safety observations, and group comparisons were only made for the data relevant to this variable.

Certain guidelines for performing a fine-grained analysis of behavioral data have been developed and are widely accepted (Mawhinney & Austin, 1999; Parsonson, 1999; Parsonson & Baer, 1992). According to Parsonson, the fine-grained analysis of graphs relies upon inspection and evaluation of several characteristics.

In particular, the following data path characteristics systematically are examined: Changes in level, trend, variability or stability and/or patterns or sequences in the data within and/or between experimental conditions; the range and any overlapping of the data points between experimental conditions; and, the number of data points in an experimental condition. (p. 47)

A fine-grained visual analysis of Figure 1, using the guidelines listed above, does not seem to reveal any substantial differences between group safety performances. In other words, there does not appear to be any substantial differences in trend, variability or stability, patterns or sequences in the data within and/or between experimental conditions between talk-aloud and silent group safety performance. Therefore, the results revealed through visual

analysis of group safety performance suggest Experiment 1 results satisfy the first silent dog method control.

Silent Dog Method Control 2

The second silent dog method control states that it must be demonstrated that task performance is altered when participants are presented with a distracter task (i.e., performance is altered because the distracter prevents self-generation of rules or self-evaluative verbalizations). Experiment 1 results clearly show safety performance, for all participants, was substantially decreased with the introduction of the distracter task, thus meeting the silent dog method control 2 requirement. The silent dog method states it is necessary to present the distracter task to talk-aloud participants, but does not specify whether or not it is necessary to present the distracter task to the participants in the silent group. Despite this, both groups of participants, in Experiment 1, were exposed to the distracter task phase. The rationale for this decision underlies the purpose, or hypothesis, of the present research. The purpose of this research is to suggest that a functional relationship exists between safety-related verbalizations and increases in safety; therefore, it seems reasonable to suggest that silent group participants also made safety-related verbalizations because of the observed safety performance improvements. As a result of this assumption, the only difference between groups was that one group made overt verbalizations (talk-aloud group) while the other possibly made covert verbalizations (silent group). Exposing both groups to the distracter phase also allowed for more replications of the distraction effect on safety performance. The substantial decreases in safety performance, which resulted from the presentation of the distracter task, clearly seem to meet the silent dog method control 2 requirements, and help support the hypothesis of the present research.

Silent Dog Method Control 3

One of the main purposes of Experiment 1 was to help establish the framework for the methods used in Experiment 2 (control 3 requirement). According to Hayes et al. (1998), the third control ensures that the absence of an effect between performance in the talk-aloud

and silent condition (the results of control 1, Experiment 1) is not because the task is contingency-shaped or that verbalizations are automatic or task irrelevant. This is achieved by employing relevant elements of the verbal reports (produced in control 1) as an external rule to new participants in Experiment 2. Because it was concluded that there was no practical difference between talk-aloud and silent group safety performance during the observation phase, thus meeting the requirements of the first silent dog method control, the verbal reports produced by Experiment 1 talk-aloud participants were used to generate description of safety rules which were presented to Experiment 2 participants in place of the independent variable: conducting safety observations.

EXPERIMENT 2

METHOD

Participants and Setting

Participants were four undergraduate students, between 18 and 22 years of age, at a midwestern university. The study took place in the same research lab as Experiment 1, and all participants were compensated \$5.00 an hour for their participation throughout the study and were paid at the end of every session.

Definition of Dependent Variables

Safety performance. The same safety performance definitions described in Experiment 1 were applied to Experiment 2.

Rule descriptions and safety-related verbalizations. A rule description statement was defined as the complete, or partial, repetition of the safety rules presented to participants during the description of rules phase. A safety-related verbalization was defined as any vocal-verbalization that was related to safety, specifically, statements related to any of the three target behaviors (i.e., back, shoulder or feet position) exposed to the rule description intervention. The total number of verbalizations which were categorized as either safety rule description statements or safety-related verbalizations was calculated and the percentage of verbalizations which were categorized into each of the two categories was calculated.

Procedures

Talk-aloud training. All participants were trained on how to talk or “think” aloud using the same training procedures followed as those for Experiment 1. Participants were instructed to talk aloud during all phases of Experiment 2.

Baseline. At the start of each baseline session, all participants were asked to perform the assembly task described in Experiment 1, and continuously speak aloud during the session.

Rule description phase. The rule descriptions were developed based on the verbal data analysis from the first experiment. These “rules” were generated using the segments that were coded as “general safety-related verbalizations.” The reason these rules were “general” versus “specific” is because the general safety-related verbalizations occurred with much more frequency than did the more specific safety-related statements. Of the safety-related verbalizations that occurred during Experiment 1, 89.4% of them were coded as “general safety-related,” and 10.6% were coded as “specific safety-related.” Therefore, we chose to model the “rules” for this phase after the general safety statements. The “general rule” used to describe safe feet position, *I need to keep my feet flat on the floor. I should keep my back and shoulders straight and not slouch or lean in any direction*, was used to describe safe back and shoulder position.

At the start of each session during this phase, participants were given descriptions of safety rules for one, or two, of the three target behaviors. Two participants comprised “Group A” and were exposed to rules for one target behavior (feet position when sitting). The remaining two participants comprised “Group B” and were exposed to rules for the other two target behaviors (back and shoulder position when sitting).

At the start of each session during the rule description phase, and before going into the work environment, participants were asked to review the rule descriptions until they memorized the rules. Participants were “quizzed” on the rules before each session throughout the phase by the researcher to ensure that participants had learned the rules, and then participants were instructed to prompt themselves on these rules at the start of each work session and anytime they noticed a change in the

position of the behaviors described by the rules. The procedures then imitated those during baseline sessions. After performance stabilized for the first target behavior(s) introduced to the rule description phase, the remaining target behavior(s) were added to the rule descriptions.

Verbal Data Analysis

The videotapes used to collect safety performance data were used to obtain rule description and safety-related verbalization occurrence data. Two experimenters collected both the above-mentioned safety performance data and data on the occurrences of rule description statements and safety-related verbalizations. The procedures used to collect these data were the same as those applied to Experiment 1. The percentage of sessions with rule description statements and/or safety-related verbalizations and the number of occurrences within each session were calculated at the end of the study. Each rule description statement and safety-related verbalization was transcribed after each description of rules phase session throughout Experiment 2.

Interobserver Agreement

IOA was conducted on all safety performance and averaged 99.8%. IOA was also conducted on (a) the transcriptions of all rule description statements and safety-related verbalizations and on (b) the coding of verbalizations as “safety rule descriptions” or “variations of the safety rule.” Two experimenters compared the written text with the audio statements and compared the coded verbalizations. IOA was calculated by dividing the total number of agreements by the total number of agreements plus disagreements and multiplying by 100%. IOA on both sets of data averaged 100%.

Experimental Design

A multiple baseline across behaviors design was used to evaluate the impact of the rule statements on safety behavior. The rule statement was first implemented for one (or two) of the three target behaviors. After performance on the first behavior(s) stabilized, the next behavior(s) were exposed to the rules.

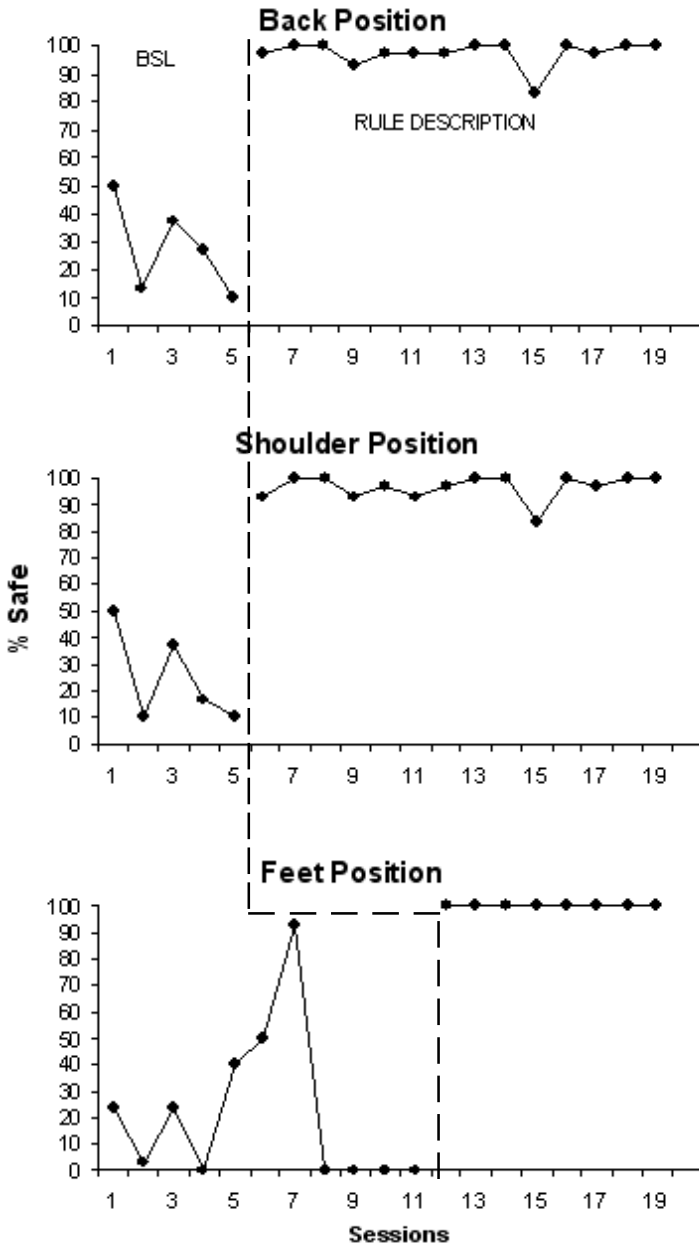


Fig. 2. Safety performance data for participant 12B.

RESULTS

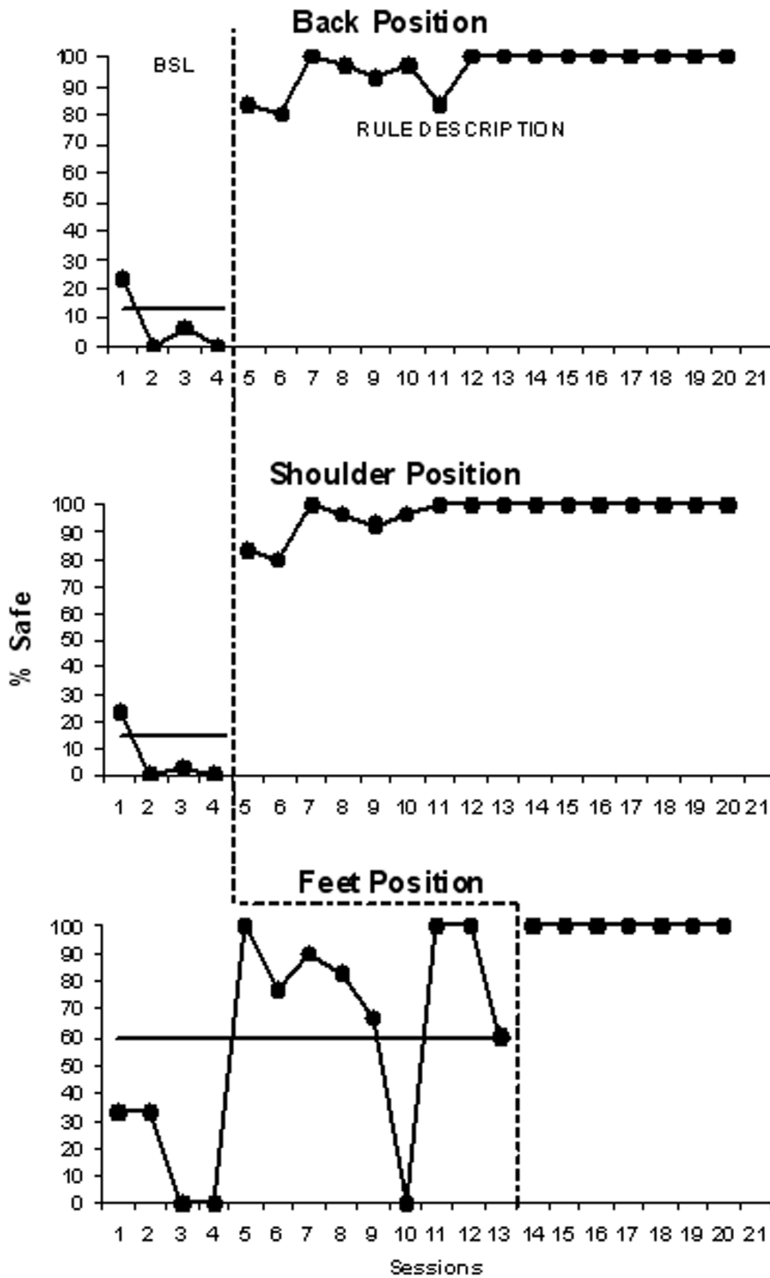
Safety Performance

Figures 2 through 5 display the safety performance for participants 12B through 15A, respectively. Each figure shows safety performance for each dependent variable during each

session and across baseline and the rule description phase.

Rule Descriptions and Safety-related Verbalizations

All four participants made safety rule description statements and/or safety-related ver-



Participant 13B

Fig. 3. Safety performance data for participant 13B.

balizations during the description of rules phase. The percentage of rule description phase sessions with at least one safety rule description statement or safety-related verbalization was 100% for participant 12B, 38% for participant 13B, and 100% for both participants 14A and 15A. Across the four participants, the

average number of safety statements (both rules and variations thereof) was 1.08 per session during the rule description phase. Out of a total of 81 safety statements, the percentage of verbalizations categorized as rule description statements was 59, and 41% were categorized as safety-related statements.

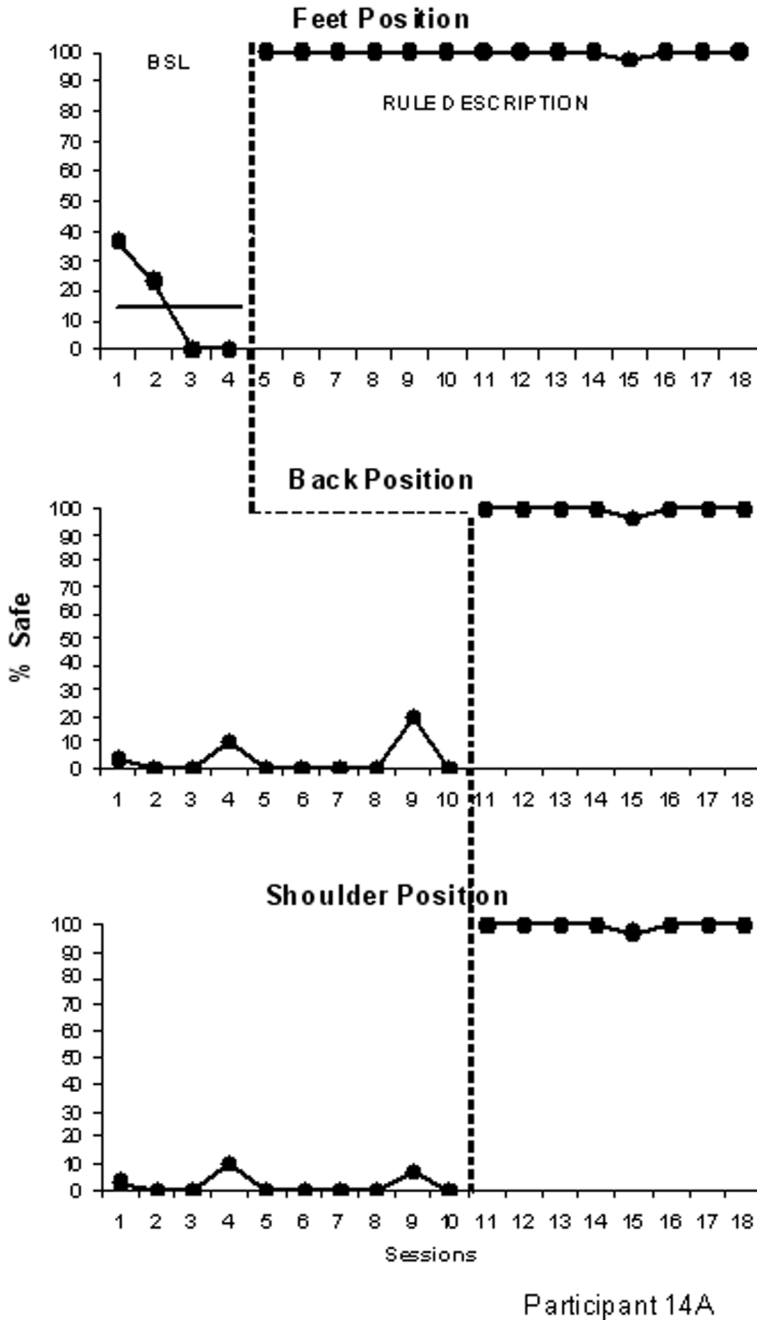


Fig. 4. Safety performance data for participant 14A.

DISCUSSION

Table 2 summarizes the total number of safety statements made by each participant during the rules phase, and the percentage of these statements categorized as complete repetition of the safety rules, or safety rule description statements.

The safety verbalizations made by Experiment 2 participants raise some interesting issues not commonly encountered in experiments that typically employ protocol analysis or the silent dog method. Protocol analysis is most

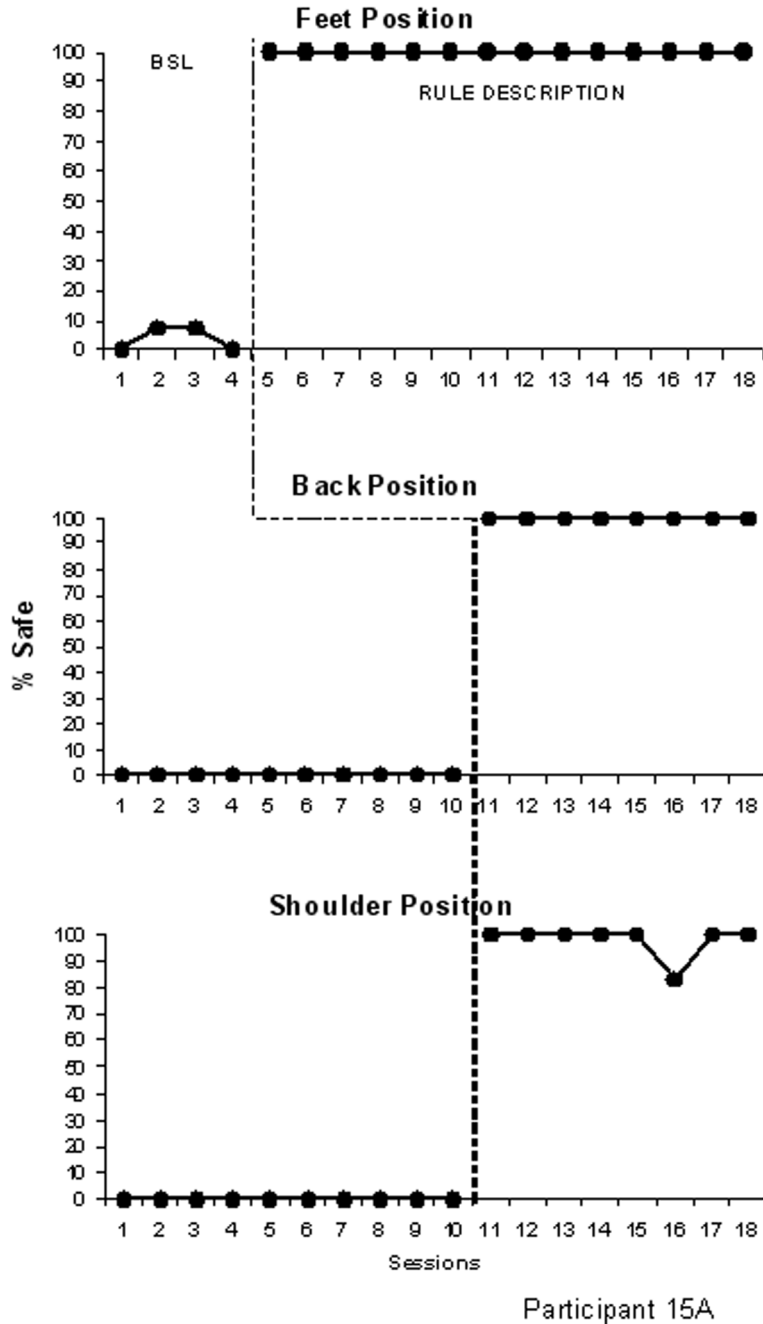


Fig. 5. Safety performance data for participant 15A.

often used for problem-solving tasks or tasks that are well-defined and have only one correct answer (Austin & Delaney, 1998; Ericsson & Simon, 1993). Unlike these traditional problem-solving tasks, the present task of interest, safety performance, is not well-defined in terms of having only one correct answer, and there-

fore, task-related verbalizations did not fall into a pattern of common actions that frequently occur during problem-solving (e.g., Bhaskar & Simon, 1977). The safety-related verbalizations made by Experiment 1 participants varied, and as a result were coded into one of two “broad” descriptor categories: general safety-

Table 2
Distribution of safety verbalizations across rules and description of rules categories.

Participant	# of Statements	% Rules	% Description
12B	23	100	—
13B	8*	25	75
14A	32	37.5	62.5
15A	18	61	39

related and specific-safety-related verbalizations. These various general safety-related verbalizations were used to generate the descriptions of safety rules presented to Experiment 2 participants. Although safety performance increased substantially for all Experiment 2 participants with the presentation of the descriptions of safety rules, what may be of most interest were the safety-related statements made by the participants. Three of the four participants made novel safety-related verbalizations, defined as variations from the descriptions of safety rules presented to them during the corresponding phase. These variations made up 41% of all safety statements made during Experiment 2.

The reason these verbalizations may be of such interest relates back to the issue regarding the “typical” studies which employ the silent dog method. This method is most commonly used to determine if a functional relationship exists between “specific” verbalizations and task performance. Because Experiment 2 resulted in variations of safety statements, rather than repetitions of the descriptions of safety rules, it may be possible to generalize conclusions regarding the functional relationship between verbalizations and performance to a broad category of safety verbalizations, versus particular words.

At this point, it important to recall the purpose of silent dog method control 3 was to demonstrate that safety performance was rule-governed and safety-related verbalizations were task relevant. According to Hayes et al. (1998) this could be achieved by presenting the safety-related verbalizations (made by Experiment 1 participants) as external rules to Experiment 2 participants in an attempt to alter, or increase, safety performance. Although Experiment 2 participants did not always repeat the description of safety rules verbatim, their safety performance did improve as a result of the pre-

sentation of the descriptions. Therefore, the results of Experiment 2 suggest the third control requirements described in the silent dog method were met.

GENERAL DISCUSSION

The overall purpose of this research was to help determine the behavioral function of conducting safety observations. We hypothesized that safety observers make safety-related verbalizations as a result of conducting observations. Therefore, the more specific objectives of this research were to determine if, in fact, safety observers make these verbalizations, and if so, to determine if the safety-related verbalizations were functionally related to increases in safety performance.

The results from both Experiments 1 and 2 appear to satisfy all three of the silent dog method control requirements, and thus, provide strong support for the existence of a functional relationship between safety-related verbalizations and increases in safety performance. These results also seem to support the suggestion that conducting safety observations may serve (a) a rule generating function, and/or (b) an antecedent, or self-monitoring, function. This is a first step toward helping determine the behavioral function of conducting safety observations. Discussing (a) how the results of this research support the above-mentioned behavioral functions, (b) the strengths and limitations of this study, and (c) suggesting future research will place us even closer toward understanding the observation process employed by behavioral safety processes.

Rule Generating Function

The results of this research suggest participants established some type of a rule as a result of conducting observations. Several par-

ticipants made safety-related verbalizations, after conducting safety observations, that seem to indicate they were following a rule. Examples of these statements include, "I want to put my arms on the armrests, but then I'll be leaning," "Can't lean back in the chair," "I need to remember to keep my feet flat on the floor," "I should keep my back straight and have good posture," and "oops, that's two to three minutes of being unsafe." According to the analysis provided by Malott (1992), it is plausible that participants stated rules describing indirect-acting contingencies (e.g., "I should perform safely so I can be scored safe," or "I should perform safely so I can please the researcher"), and safe behaviors were controlled by direct-acting escape contingencies (e.g., fear of not being scored safe, or fear of displeasing the researcher). The verbal reports measured throughout this study did not yield any complete contingency-specifying rule statements. In other words, participants did not explicitly describe aloud the indirect-acting-contingencies they may have generated after conducting safety observations. Some researchers suggest it may not be necessary for verbal stimuli to overtly include contingency-specifying stimuli in order to be considered a rule (Blakely & Schlinger, 1987; Schlinger, 1993). According to Schlinger,

In sophisticated listeners it is common for a single word to have function-altering effects. For example, if someone is about to eat some food and someone else says, "poison," this alters the evocative function of the sight of the food in the sense that it decreases the momentary probability of the behavior of putting it in the mouth. Some might suggest that there is an implicit *contingency-specifying stimulus* (CSS) in the word "poison" (e.g., "If you eat this food you will get very sick"). (p. 12)

Therefore, regardless of the form of the verbal stimulus, if it alters the function of a stimulus then it may be considered a rule (Schlinger). Hence, it is plausible that some of the statements observed throughout this experiment can be considered rules, and the above-quoted statement perhaps explains why explicit contingency-specifying statements were not observed. Regardless of this issue, the verbal reports do seem to provide enough evidence to suggest that participants self-generated rules as a result of conducting observations.

Self-Monitoring

The verbal reports made by Experiment 1 participants indicate they self-monitored safety performance after conducting safety observations. Some examples of safety-related verbalizations that support this theory include: "I almost moved my feet," "Feet on the floor," "I'm slouching," and "Sitting properly, sitting properly." An often cited explanation for the effectiveness of self-monitoring as an independent variable is the "reactive effect" (Nelson et al., 1982), and this explanation may also describe the effects of conducting safety observations. Several widely accepted views explaining reactivity exist (Kanfer & Gaelick-Buys, 1991; Rachlin, 1974; Hayes & Nelson, 1983), and all of them attribute the changes in performance to a different component of the self-recording procedure. Kanfer and Gaelick-Buys might suggest that a participant would self-compare his or her safety behaviors to the definitions provided on the checklist, or with confederate performance, then he or she would self-deliver consequences contingent on his or her own performance of the target behaviors. Rachlin might propose that the recording response, the self-administered consequences, or a combination of the two serve as cues to "remind" the participant of the external environmental consequences that actually control response frequency. Hayes and Nelson might add that the entire self-recording procedure (e.g., the video, the checklist, the target behaviors, the confederate's performance) serves as an initiator of reactivity. In other words, everything associated with the observation process would make more obvious the environmental consequences, or perhaps self-generated consequences (such as those specified in self-generated rule statements).

Self-Generation of Rules vs. Self-Monitoring?

Several questions arise as a result of discussing the possible rule-generating and self-monitoring functions of safety observations. First is whether it is even logical to separate the two effects. How can a person ensure he or she is complying with a rule if he or she does not monitor behavior? A person may engage in safe behaviors to escape a self-generated aversive condition, such as noncompliance to a safety

rule, but in order to perform safely he or she must self-monitor relevant safety behavior. Another interesting issue concerns the temporal relationship between the two effects (assuming one is not exclusive of the other). Did participants begin to self-monitor safety performance as a result of conducting observations which then resulted in the self-generation of safety rules? Or, did participants self-generate rules as a result of conducting observations, which then resulted in self-monitoring of safety performance? Unfortunately, the results of this research cannot provide concrete answers to these questions, and may actually result in more questions.

The intervention to which Experiment 2 participants were exposed involved two components: rules and a "suggestion" to self-monitor. Participants were (a) provided with descriptions of safety rules and (b) were instructed to state these descriptions at the start of the session and anytime they noticed a change in the relevant behavior(s). Although participants were not explicitly instructed to follow the rules, just to repeat them, it is plausible that the latter part of these instructions (part b) "suggested" participants should follow the rules and self-monitor their performance. The results of Experiment 2 indicate participants followed the descriptions of safety rules (safety performance substantially increased during the corresponding phase) and self-monitored their safety performance (repetition of descriptions of safety rules and novel safety-verbalizations were made). Therefore, it is not known (a) whether participants would have self-monitored their performance had they only been provided with the rules (and not asked to repeat them), or (b) whether participants would have self-generated rules as a result of being instructed to self-monitor but not given the descriptions of safety rules.

Strengths and Limitations

One of the most notable strengths of the present study is the use, and thorough analysis, of verbal data to help determine the behavioral function of an intervention (conducting safety observations). The use of verbal reports to assess their effects on performance, and help determine the behavioral function of interventions used in the field of organizational behavior management has been suggested many times. Agnew and Redmon (1992) suggested

the "analyses of contingencies based on rules raises several possibilities for future OBM research" (p. 72), and suggested participants should be interviewed after an experiment to determine the nature of rules used. Perone (1988) suggested that more use should be made of self-report data, and also suggested verbal reports be used (a) as primary data, and (b) for explanatory purposes in understanding behavior patterns. Mawhinney and Austin (1999) suggested the use of concurrent verbal reports could help explain differences in levels of performance. Despite the prevalence of these suggestions, few in the field of OBM, have implemented them (Agnew & Redmon, 1992).

The use of verbal reports was a strength of the study, and the results of this research are the first to provide support for the existence of a functional relationship between safety-related verbalizations and safety performance. These results have also brought us one step closer toward understanding the behavioral mechanisms responsible for the effectiveness of conducting safety observations, and thus, may be able to help improve the effectiveness of the behavioral safety process.

Another strength of this study is that it was a first attempt to help determine the behavioral function of conducting safety observations. Although conducting observations has been demonstrated to be an effective intervention for increasing safety performance (Alvero & Austin, 2003, 2004; Sasson, 2002), this was the first attempt to help determine the behavioral mechanisms responsible for its effectiveness. The results of this first attempt can serve as a framework for future research, which may result in a more "concrete" answer to the present experimental question.

A unique strength of this study is the fact that it is the first to use protocol analysis and the silent dog method to measure and evaluate the verbalizations relevant to an ill-defined task. As previously mentioned, the experiments which typically employ protocol analysis and the silent dog method, involve tasks that are well-defined and have only one answer (Ball, Langholtz, Auble, & Sopchak, 1998; Bhaskar & Simon, 1977; Wulfert, Dougher, & Greenway, 1991). Unlike these typical experiments, the present experiment attempted to record and analyze verbalizations related to safety performance: a task that occurs concurrently with many other tasks, and is not well-

defined in terms of having only one answer. This unique strength could also be considered a limitation. The ill-defined nature of safety performance resulted in a large quantity of “irrelevant” verbal data (98.8% of segments coded), and a very small quantity of safety-relevant data (1.2% of segments coded). The high percentage of “irrelevant” data are not very surprising when one considers the nature of the assembly task. Participants were required to thread beads onto a string in a specific color sequence. As a result, the majority of the verbalizations were directly related to this task (e.g., “blue, red, green, orange, blue”), and only 1.2% of the verbalizations made by Experiment 1 participants were related to the “subtask” of safety performance. Therefore, the small percentage of safety-related verbalizations were the only data used to help determine if a functional relationship existed between safety-related verbalizations and increases in safety performance. Another limitation related to this issue were the omissions of safety-related verbalizations from several participant verbal reports (participants 1A and 3B), and from all participant verbal reports during the observation reversal phase. The talk-aloud training procedures used throughout this study were largely based on the procedures employed by protocol analysis (Ericsson & Simon, 1993), and thus, were designed to collect verbal reports from well-defined tasks that are well-defined and only have one answer. Therefore, these training procedures may not be well-equipped to capture verbalizations relevant to an ill-defined task, such as safety performance. The answers given by participant 1A during the exit interview seem to support this conclusion. When informed that she had not made any verbalizations regarding safety performance, and then asked if this information surprised her, participant 1A replied, “yes, I thought I did (make safety verbalizations) sometimes.” Participant 1A was then asked why she did not make these verbalizations, and she answered, “I think because I thought about it so quickly and while I was doing something else, and a lot of times it was right in the beginning, so it never crossed my mind to say it out loud because it was so fast.” It is more difficult to speculate regarding the omission of safety-related verbalizations during the observation reversal phase. One possible explanation is that participants simply “forgot” the talk-

aloud training procedures after exposure to the distracter task and were eager to complete the study.

Future Research

Future research endeavors designed to determine the behavioral function of conducting safety observations should consider the above-mentioned limitations and modify procedures accordingly. For example, the omission of explicit contingency-specifying rule statements may have been prevented by having instructed participants to talk aloud throughout the entire observation procedure, before they began the assembly bead task, and as they performed the bead task. It is plausible that participants self-generated explicit rule statements while conducting safety observations or as they prepared to begin the assembly task. Advancing the start of the talk-aloud procedure to include the observation process may “capture” these statements, assuming such statements were made. Unfortunately, whether such contingency-specifying rule statements were made cannot be determined based on the results of the present research. Another suggestion, related to this issue, concerns the talk-aloud training procedures used in the present research, which involved solving arithmetic problems aloud. As previously mentioned, these training procedures were based on the procedures employed by protocol analysis (Ericsson & Simon, 1993), and thus, were designed to collect verbal reports from “primary” tasks versus ill-defined tasks, such as safety performance. Increasing the similarity between the training procedure condition and the experimental condition may help increase the probability of recording verbalizations related to all tasks performed by the participants.

In summary, the present study involved exhaustive data collection procedures and required time-consuming data analyses, but the results are significant and are considered worth the effort. The results of this study (a) are the first to demonstrate the existence of a functional relationship between safety-related verbalizations and increases in safety performance, (b) provide a framework for similar future endeavors in the field of OBM, specifically in the area of safety, and (c) raise interesting questions concerning the possible behavioral functions of conducting safety observations.

REFERENCES

- Agnew, J. L., & Redmon, W. K. (1992). Contingency specifying stimuli: The role of "rules" in organizational behavior management. *Journal of Organizational Behavior Management, 12*(2), 67–74.
- Alvero, A. M. & Austin, J. (2003). The observer effect. In T. McSween (Ed.), *The values-based safety process* (2nd ed.). New York, NY: John Wiley & Sons, Inc.
- Alvero, A. M., & Austin, J. (2004). The observer effect: The effects of conducting behavioral observations on the behavior of the observer. *Journal of Applied Behavior Analysis, 37*, 457–468.
- Austin, J., & Delaney, P. F. (1998). Protocol analysis as a tool for behavior analysis. *The Analysis of Verbal Behavior, 15*, 41–56.
- Baer, D. M. (1977). Perhaps it would be better not to know everything. *Journal of Applied Behavior Analysis, 10*, 167–172.
- Ball, C. T., Langholtz, H. J., Auble, J., & Sopchak, B. (1998). Resource-allocation strategies: A verbal protocol analysis. *Organizational Behavior and Human Decision Processes, 76*, 70–88.
- Bhaskar, R., & Simon, H. A. (1977). Problem solving in semantically rich domains: An example from engineering thermodynamics. *Cognitive Science, 1*, 193–215.
- Blakely, E., & Schlinger, H. (1987). Rules: Function-altering contingency-specifying stimuli. *The Behavior Analyst, 10*, 183–187.
- Doyle, A. C. (1892). The memoirs of Sherlock Holmes. Reprinted in *Sherlock Holmes: Complete novels and stories* (Vol. 1, pp. 455–477). New York: Bantam Books, 1986.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. (rev. ed.). Cambridge, MA: MIT Press.
- Hayes, S. C., & Nelson, R. O. (1983). Similar reactivity produced by external cues and self-monitoring. *Behavior Modification, 7*, 183–196.
- Hayes, S. C., White, D., & Bissett, R. T. (1998). Protocol analysis and the "silent dog" method of analyzing the impact of self-generated rules. *The Analysis of Verbal Behavior, 15*, 57–63.
- Huitema, B. E. (1985). Autocorrelation in applied behavior analysis: A myth. *Behavioral Assessment, 7*, 107–118.
- Kanfer, F. H., & Gaelick-Buys, L. (1991). Self-management methods. In F. H. Kanfer & A. P. Goldstein (Eds.), *Helping people change* (4th ed.) (pp. 305–360). New York: Pergamon.
- Komaki, J. (1986). Toward effective supervision: An operant analysis and comparison of managers at work. *Journal of Applied Psychology, 71*, 270–279.
- Komaki, J., Barwick, K. K., & 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., Heinzmann, A. T., & Lawson, L. (1980). Effect of training and feedback: Component analysis of a behavioral safety program. *Journal of Applied Psychology, 65*, 261–270.
- Krause, T. R. (1997). *The behavior-based safety process*. New York: Van Nostrand Reinhold.
- Malott, R. W. (1992). A theory of rule-governed behavior and organizational behavior management. *Journal of Organizational Behavior Management, 12*(2), 45–65.
- Mawhinney, T. C., & Austin, J. (1999). Using concurrent verbal reports to examine data analyst verbal behavior. *Journal of Organizational Behavior Management, 18*(4), 61–81.
- McSween, T. E. (2003). *The values-based safety process* (2nd ed.). New York: John Wiley & Sons, Inc.
- National Institute for Occupational Safety and Health. (1998). *OSHA Ergonomics Report DT931018* [Data file]. Available from: <http://www.osha-slc.gov/SLTC/ergonomics/ergo-hiosh.html>.
- Office of Health and Safety Information System. (1998). *Safety manuals* [Data file]. Available from: <http://www.cdc.gov/od/ohs/default.htm>.
- Parsonson, B. S. (1999). Evaluating experimental and quasi-experimental field data. *Journal of Organizational Behavior Management, 18*(4), 47–51.
- Parsonson, B. S., & Baer, D. M. (1992). The visual analysis of data, and current research into the stimuli controlling it. In T. R. Kratochwill, & J. R. Levin (Eds.), *Single-case research design and analysis: New directions for psychology and education* (pp. 15–40). Hillsdale, NJ: Erlbaum Associates, Inc.

- Perone, M. (1988). Laboratory lore and research practices in the experimental analysis of human behavior: Use and abuse of subjects' verbal reports. *The Behavior Analyst, 11*, 71–75.
- Rachlin, H. (1974). Self-control. *Behaviorism, 2*, 94–107.
- Sasson, J. (2002). *Examining the effects of conducting behavior-based safety observations*. Unpublished master's thesis, Western Michigan University, Kalamazoo.
- Schlinger, H. D. (1993). Separating discriminative and function-altering effects of verbal stimuli. *The Behavior Analyst, 16*, 9–23.
- Sulzer-Azaroff, B., & Fellner, D. (1984). Searching for performance targets in the behavioral analysis of occupational health and safety: An assessment strategy. *Journal of Organizational Behavior Management, 6*(2), 53–65.
- Wulfert, E., Dougher, M. J., & Greenway, D. E. (1991). Protocol analysis of the correspondence of verbal behavior and equivalence class formation. *Journal of the Experimental Analysis of Behavior, 56*, 489–504.