Journal of Safety Research, Vol. 15, pp. 125–135, 1984

0022-4375/84/ $3.00 + .00
Printed in the USA

Measuring the Effectiveness of an Industrial Lift Truck Safety Training Program

H. Harvey Cohen and Roger C. Jensen

A behavior (work) sampling approach was used to both develop and evaluate the effectiveness of an occupational safety training program for industrial lift truck operators. Two studies, each using different experimental designs and performed at two separate warehouses, were conducted, resulting in a total of 96 operators trained. Observations through several months of training program assessment indicated that: (a) occupational safety training, emphasizing modification of operationally defined unsafe work practices derived from task/hazard analysis, can be demonstrated to be effective and to endure beyond cessation of performance feedback; (b) the basis for endurance appears to be continued practice in the modified safe work procedures, coupled with a redefinition of group norms sustained through informal influences such as peer modeling of desired behaviors and continued management support of the program; and (c) a behavior sampling procedure, specifying performance-based criteria, can be used effectively in both the development and evaluation of an occupational safety training program.

Over the years, safety training has been heavily depended upon as a means of reducing occupational accidents, yet the literature lacks reports of definitive research demonstrating the value of safety training and the length of its effectiveness (Surry, 1969). The literature that exists consists largely of course descriptions, lesson plans, and programs in use by companies or proposed by individuals. Few of these reports are based on an assessment of need through task analysis, and only one study has been found that measured the effects of safety training on on-the-job performance (Komaki, Heinzmann, & Lawson, 1980).

The actions of management following training are critical in affecting on-the-job performance. Post-training actions can range from no follow-up to extensive programs involving (a) goal-setting, (b) performance monitoring, (c) feedback, and (d) rewards.

H. Harvey Cohen, PhD, is Co-owner and Principal Scientist of SAFETY SCIENCES, Inc., San Diego, California 92121. Roger C. Jensen, JD, is currently on academic leave from his position as Chief of the Accident and Injury Epidemiology Branch, Division of Safety Research, National Institute for Occupational Safety and Health (NIOSH), while pursuing a PhD.

This study was performed by Safety Sciences under contract 210-79-0018 with NIOSH. The authors wish to acknowledge the valuable technical contributions made by Margaret Hubbard Jones.

Goal-setting is an approach for improving performance on tasks in which performance is largely a function of the workers' conscious aspirations (Das, 1982). The goals may be established by management or by a worker participation approach. To be effective the goals should be accepted and achievable by the individual or group for which they are established (Locke, 1968). *Feedback* is a means of improving or sustaining performance by providing information about actual performance to the worker (Das, 1982; McCormick & Sanders, 1982; Meister, 1976). *Rewards* are used to reinforce desired behaviors and to provide motivation for sustaining achievable performance levels (Cohen, Smith, & Anger, 1979). All of these approaches require an objective and meaningful measure of performance.

An objective measure of performance is also essential when conducting research into the effectiveness of training and other management programs for improving worker safety. Traditional measures of safety performance, such as lost-time accidents, are "rare events" in the statistical sense and, consequently, not sensitive enough to evaluate the effectiveness of specific intervention programs in a single establishment.

A more sensitive measure (Rockwell, 1959; Tarrants, 1980) uses operationally defined performance criteria for measuring the effectiveness of program intervention through a work sampling procedure. This procedure is commonly used by industrial engineers for making determinations such as the portion of time a particular machine is in use. When work sampling is used for monitoring human behavior, the behavior may be dichotomized into categories such as proper or improper, wearing or not wearing ear protection, standing or walking, etc.

Rockwell indicated that such surrogate safety performance measures should be: (a) *observable* in order that they can be measured; (b) *quantifiable* in order to permit the use of statistical inference; (c) *reliable* to the extent that they provide minimum variability when measuring the same condition; and (d) *valid* in that they are related to factors that precipitate accidents.

A few studies have used behavioral sampling methods to evaluate the effects of safety training programs in various work settings (Komaki, Barwick, & Scott, 1978; Komaki et al., 1980; Smith, Anger, & Uslan, 1978; Zohar, Cohen, & Azar, 1980). The study by Komaki, Barwick, & Scott (1978) used a behavioral sampling approach to demonstrate the beneficial effects of a combined safety training plus performance feedback intervention program. This was followed by a second workplace study (Komaki et al., 1980) designed to determine the effects of safety training separately from the effects of performance feedback. The study found that safety training alone resulted in improved performance, while training combined with post-training feedback yielded even better performance.

The extent to which employee training and/or management programs produce improved behavior following cessation of the formal intervention program is a related question that needs investigation. Komaki et al. (1978) found that when the research team stopped monitoring the workers' performance and providing feedback, performance returned to the pre-intervention level. Zohar (1980) suggests that a "holistic modification approach" in which workers' behavior and managerial standards are concurrently modified may be necessary to sustain changes in group norms. In such an interactive system, new workers introduced to the plant would be encouraged by both management and their peers through ongoing performance feedback to model the modified group norms, thereby sustaining the desired group performance.

The two studies described in this paper were initiated by the National Institute for Occupational Safety and Health (NIOSH) in order to assess the value of occupational safety training and the influence of post-training management actions on the safety performance of workers in a semi-skilled job. Industrial lift truck operators were selected for study because: (a) industrial lift trucks are widely used throughout industry; (b) lift truck operations have demonstrated a high risk and propensity for accidents; and (c) the risk of injury associated with typical lift truck operations depends to a large extent on the performance of the operators.

METHOD

The basic approach followed Goldstein's model (1974) which specifies three major phases for a quality training program: (a) needs assessment, (b) program development, and (c) program evaluation.

Needs Assessment

The needs assessment involved the determination of existing knowledge concerning lift truck safety and an analysis of injury data. The search for existing knowledge identified: (a) research and technical reports related to lift truck safety and behavioral observation methodology; (b) ANSI (American National Standards Institute) and OSHA standards for safe lift truck operation; (c) operating manuals from a variety of lift truck manufacturers; and (d) all available media, both movie and slide presentations, and workbook training courses on industrial lift truck operation and safety. None of the training materials was found to be suitable for the purposes of these studies.

The injury data analysis consisted of: (a) approximately 1,000 lift truck accident reports from a national data base of over 10,000 general industry accident cases and (b) several hundred lift truck accident reports from the two warehouse participants over a period of 3 years preceding the studies.

Training Program Development

The program development involved a series of tasks. Pre-baseline observations of both warehouse sites were made in order to identify those recurrent behaviors most suitable for training as well as to become familiar with the warehouse facilities.

A detailed task/hazard analysis was conducted in order to derive a rational basis for developing the training program. Based on the above observations, each task was broken down by: (a) the necessary knowledge and skill requirements, (b) the potential consequences of behavioral errors, and (c) criticality ratings based on the injury data analysis.

From the task/hazard analysis and the pre-baseline observations of warehouse operations, operator behaviors were identified that were: (a) capable of being operationally defined, i.e., measurable; (b) frequently observable; (c) capable of being reliably observed; (d) related to frequent accident occurrence, i.e., valid; and (e) amenable to corrective action through training (rather than through equipment or job redesign).

The training materials were a slide/sound presentation and an accompanying instructor's manual. The manual presented: (a) specific training and behavioral objectives, (b) a copy of the script with pictures, and (c) detailed instructions for course implementation.

The training program consisted of five sessions: an introductory session, three instructional sessions, and a "hands-on" practice exercise. The sessions took about 20 to 45 minutes each and were delivered on five successive work days. Each of the three instructional sessions was developed around five critical behaviors that met the previously defined criteria (i.e., were measurable, observable, valid, etc.). There were thus 15 operationally defined safe work practices around which the training and performance feedback program was based.

The three instructional sessions were set up in the following manner: A first slide introduced the situation to be trained. A second slide or set of slides showed typical *incorrect* ways of handling the situation. A third slide or set of slides showed the *correct* procedure. This approach ensured that the proper safe work practice was the last bit of information presented to the trainees.

Active learning was promoted through two types of trainee participation. First, semi-structured discussion immediately followed each training item (behavior). Second, in the final session, a practice exercise involved performance of all behaviors and used a peer modeling approach to reinforce correct work practice. This approach required all trainees to observe and score the performance of each trainee as he went through the practice course layout. Individual scores were specifically *not* compared to avoid promoting competition.

A post-training management program was also developed that consisted of several elements. First, daily feedback was provided in the form of verbal and posted summaries of group performance. This was combined with group performance goal setting, i.e., a goal of better than 80% overall correct performance was agreed upon by the group. The trainer, a first-line supervisor meeting specific criteria of leadership, experience, and communication skills, provided performance feedback, including individualized coaching as necessary, in a positive, constructive, and confidential manner. Thus, progressive increments in desired behavior were shaped on an individual basis. Finally, all levels of management had input into the development and execution of both the training and management program and firmly supported it throughout. The program was designed to be a company program, that is, it met company needs, was developed and executed with management assistance, and was provided to management in final form for continued use after completion of the study.

Thus, the training was reinforced by combining: (a) trainee participation, (b) performance feedback, (c) group performance goal setting, (d) peer group modeling, and (e) management support.

Program Evaluation

The program evaluation was based on observations of on-the-job behaviors before and after training. Of the 15 criterion behaviors used in the training program, 14 proved usable as performance measures. One behavior requested by management was not included in the final analysis due to limited opportunity for observation.

Three observers were trained in precise methods for observing the criterion behaviors. The observers practiced at a practice warehouse (also the site where training program slides were taken) until their interobserver reliabilities exceeded .8 on the Cohen's Kappa Statistic (Cohen, 1960). Cohen's Kappa controls for chance agreement and is, therefore, more conservative than percent agreement. Actually, interobserver reliability exceeded .9 during the data collection phases of the study, a level considered to be exceptional. Weekly checks on observer reliability were made throughout the assessment phases of both studies.

Each warehouse was divided into eight observation locations. Observers rotated through these locations according to a schedule that ensured a random starting point and equal coverage by all observers. An observer stayed in a single location for approximately 30 minutes and then rotated stations. Observer rest breaks were interspersed throughout the daily session. All shift hours were sampled.

The 15 behaviors were listed on a data recording sheet. When a lift truck operator was observed, the observer marked a plus for each behavior that was correctly performed, a minus for each behavior that was not incorrectly performed, and a zero for each behavior that was not observed.

Frequency counts and error rates were computed daily. Computer printouts were provided to the supervisor/trainer for daily performance feedback to those operators receiving feedback. For the purpose of the studies reported in this paper, error rate was defined as:

$$\text{Error Rate} = \frac{\text{No. of Incorrect Behaviors Observed}}{\text{Total Behaviors Observed}} \times 100$$

Experimental Settings

Two studies, each employing a different experimental design, were performed. The two studies took place in two separate warehouses. The warehouses were both large regional distribution operations for two major national retailers. Both were located in Southern California and, despite different managements, displayed remarkably similar operational and employee characteristics. Both facilities were new, each employed 48 lift truck operators (all 96 of whom were eventually trained), all operators were male, and turnover was negligible in both warehouses. Table 1 presents a summary of some employee characteristics at the two warehouses.

STUDY 1

Experimental Design

The first study, conducted at Warehouse 1, used a between and within groups comparison design. Treatment groups were assigned in the following manner. First, operators were stratified by: (a) vehicle type (e.g., forks, clamps, appliance handlers, cherry

TABLE 1
SUMMARY OF LIFT TRUCK OPERATOR CHARACTERISTICS
AT THE TWO PARTICIPANT WAREHOUSES

WARE-HOUSE	AGE	MARITAL STATUS	EXPERIENCE AT JOB (YEARS)	EDUCATION LEVEL (YEARS)	ACCIDENT FREQUENCY PER OPERATOR (3 YEAR PER.)	ACCIDENT SEVERITY (LOST WORKDAYS)
1	$\bar{X} = 34.2$	Married = 71%	$\bar{X} = 7.6$	$\bar{X} = 11.7$	$\bar{X} = 1.1$	$\bar{X} = 16.7$
	$R = 23–50$	Single = 21%	$R = 1–20$	$R = 6–14$	$R = 0–4$	$R = 1–89$
		Divorced = 8%				
2	$\bar{X} = 33.9$	Married = 66%	$\bar{X} = 6.9$	$\bar{X} = 12.2$	$\bar{X} = 1.0$	$\bar{X} = 15.7$
	$R = 21–48$	Single = 25%	$R = 1–18$	$R = 8–14$	$R = 0–3$	$R = 1–76$
		Divorced = 9%				

pickers); (b) department (which corresponded to different vehicle types); and (c) relative exposure (i.e., observation frequency established during pre-baseline observations). Then, within each of the strata, all lift truck operators were randomly assigned to one of three groups (12 operators per group): (a) a training-only group, (b) a training-plus-feedback group, and (c) a control group that received no training and no feedback until after the Post-training 1 observation period. Observations were compared both before and after training separately for each of the three groups.

Another 12 employees who were occasional lift truck operators were eventually trained along with the control group at management's request. However, observational data from these occasional operators were not included in the Post-training 2 and Retention phases of the study.

The experiment was divided into four phases:

1. The *Pre-training phase* during which none of the operators had been trained;
2. The *Post-training 1 phase* during which the control group remained untrained, the treatment group had received training, and the treatment-plus-feedback group had received training and was also receiving performance feedback;
3. The *Post-training 2 phase* during which all three groups had received training but only the training-plus-feedback group received performance feedback; and
4. The *Retention phase* which started 3 months after the end of the Post-training 2 phase (and the end of the feedback program).

Observations were double-blind, i.e., at no time were either observers or operators informed that different treatments were being evaluated, nor did they know to which group operators had been assigned. The operators in the control group and the occasional operators were told that they would be trained at a later time due to schedule constraints. Operators were recognized by a reliable, two-step coding system using composite pictures and code names, the key to which was available only at the remote data processing center.

Results

Figure 1 and Table 2 show that pre-training error rates were comparable and stable for all three experimental groups, averaging .34. Following the initial training, all three groups showed a decrease in their mean error rates (Table 3) with the training-plus-feedback group showing the largest decrease (23%) followed by the training-only (18%) and control (6%) groups. The calculated F value was significant ($F\,[2, 38] = 7.58$, $p < .01$). Post-hoc analysis, using the Duncan Multiple Range Procedure, indicated that the training-only and training-plus-feedback groups each differed from the control group ($p < .05$) during the Post-training 1 phase, but did not differ significantly from one another. Toward the end of the Post-training 1 phase, the error rates of the three groups converged, suggesting that the effects of the intervention

FIGURE 1
MEAN ERROR RATES FOR SUCCESSIVE PHASES OF THE TRAINING PROGRAM (WAREHOUSE 1)

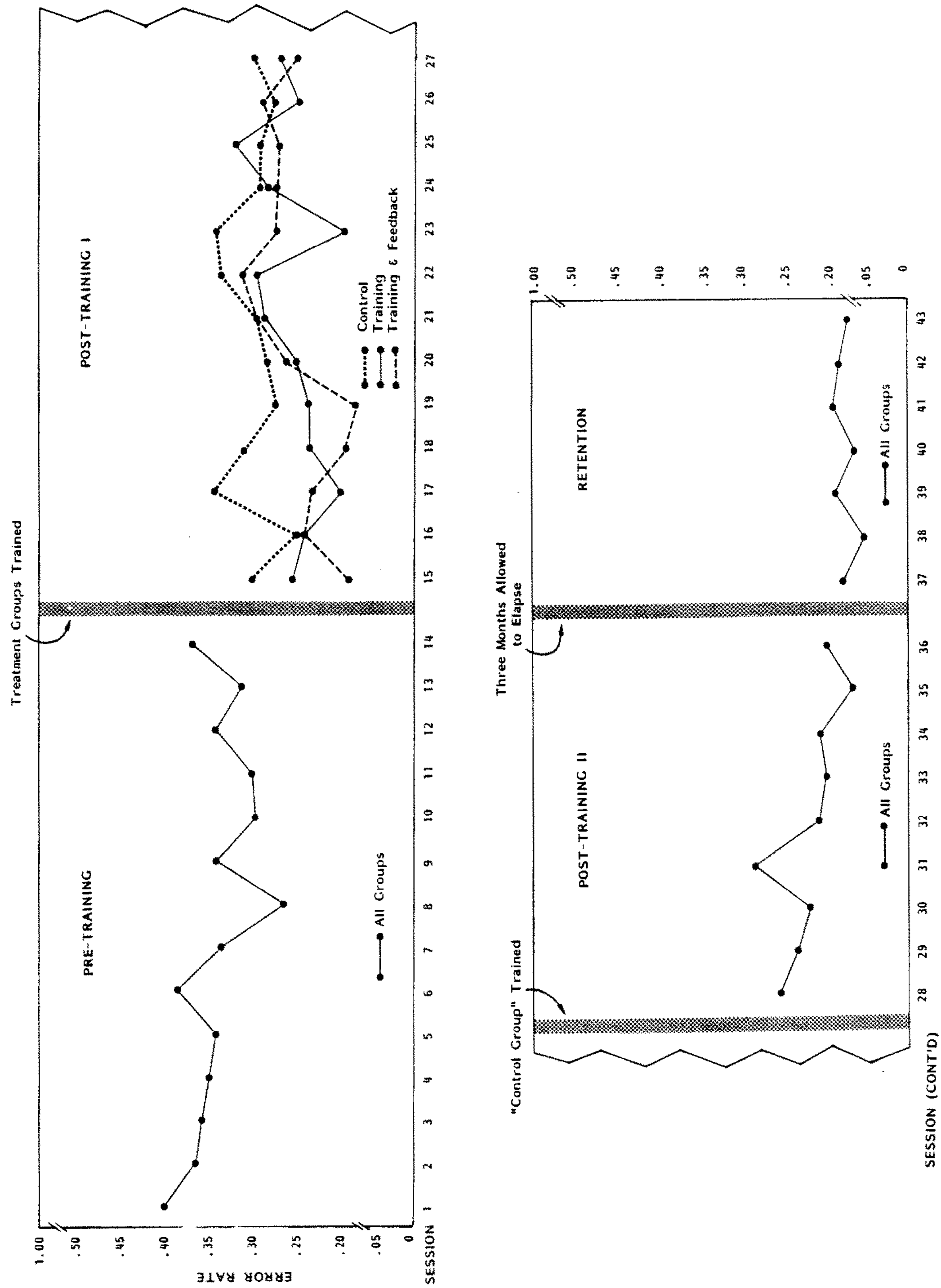

TABLE 2
SUMMARY OF MEAN ERROR RATES
(WAREHOUSE 1)

GROUP	PRE-TRAINING	POST-TRAINING I	POST-TRAINING II	RETENTION
Control	.34	.32	.23	
Training	.33	.27	.26	
Training plus feedback	.35	.27	.25	
All operators	.34	.27	.25	.19

program had begun to wear off. Observers also noted that some behaviors were becoming compromised when trained operators had to interact with untrained control group operators, particularly in conflict avoidance situations involving behaviors such as signaling and yielding at blind intersections.

Following the Post-training 1 phase the control group and occasional operators were trained so that during the Post-training 2 phase all operators had been trained. The results presented in Table 3 indicate that the performance of all three groups improved, with the greatest improvement (28%) in the original control group. Subsequent discussions with operators indicated a peer modeling influence, i.e., the control group operators were modeling the behavior of their previously trained counterparts. Similarly, behaviors in potential conflict avoidance situations were now reinforcing to the originally trained operators because nearly all operators were using the correct procedures. Further evidence of modeling may be deduced from the finding during Post-training 1 that control group operators actually showed a slight improvement of 6% over baseline *before* they were formally trained.

Each phase of training program assessment lasted about 1 month. After completion of the Post-training 2 phase, about 3 months elapsed before returning to Warehouse 1 to see if retention of safe work practices had occurred. The results shown in Figure 1 and Table 2 indicate additional improvement, with the error rate decreasing from 0.25 in the Post-training 2 phase to 0.19 in the Retention phase.

The total performance gain in Study 1 was 44% improvement from pre-training (baseline) levels. A one-way analysis of variance for performance error rate as a function of successive phases of the training program in Warehouse 1 indicated a strong treatment effect ($F\ [2, 29] = 61.67$, $p < .0001$). Post-hoc comparisons using the Duncan test demonstrated that more errors were observed during the Pre-training phase than during each of the successive phases, i.e., Post-training 1, Post-training 2, and Retention. In other words, there were significantly fewer errors at each successive phase of training program assessment.

Table 4 shows the percentage change in mean error rate for each of the 14 behaviors used for performance evaluation. Improvement occurred in 12 behaviors. Not unexpectedly, one behavior (keeps all body parts within the truck) showed no improvement from its initial low error rate of 0.01. The other behavior that did not improve was "drives in reverse." The mean error rate actually in-

TABLE 3
PERCENT DECREASE IN MEAN ERROR RATES
(WAREHOUSE 1)

GROUP	PRE-TRAINING VS. POST-TRAINING I	POST-TRAINING I VS. POST-TRAINING II	POST-TRAINING II VS. RETENTION	PRE-TRAINING VS. RETENTION
Control	6%	28%		
Training	18%	4%		
Training plus feedback	23%	7%		
All operators		11%	24%	44%

TABLE 4
PERCENT CHANGE IN ERROR RATES FOR INDIVIDUAL BEHAVIORS
(WAREHOUSE 1)

BEHAVIOR	PRE-TRAINING MEAN ERROR RATE	RETENTION MEAN ERROR RATE	PERCENT CHANGE IN ERROR RATE
1. Warns trucks	.82	.50	39%
2. Yields to trucks	.68	.44	35%
3. Warns co-workers	.89	.42	53%
4. Yields to co-workers	.93	.37	60%
5. Sounds horn at blind intersection	.86	.53	38%
6. Slows down at blind intersection	.67	.47	30%
7. Looks at blind intersection	.67	.38	43%
8. Looks in direction of travel	.33	.19	42%
9. Maintains moderate speed	.42	.20	52%
10. Avoids quick starts/ changes of direction	.25	.14	44%
11. Keeps all body parts within truck	.01	.01	0
12. Maintains forks in proper position	.31	.13	58%
13. Maintains balanced load	.13	.04	69%
14. Drives in reverse	.49	.51	−4%
All behaviors	.34	.19	44%

creased slightly from 0.49 to 0.51. Inquiry revealed that operators at Warehouse 1 resisted this change because their lift trucks were propane-powered (as opposed to clean-burning battery-powered lift trucks in Warehouse 2) and driving in reverse caused them to breathe in noxious fumes. Further, continuous looking over one's shoulder is an unnatural and uncomfortable posture to assume for prolonged periods.

Table 5 shows a breakdown of individual behaviors by number and percent of total observations. Observation rates ranged from about 1 to 16% for individual behaviors with a median of about 9%. Clearly, no single behavior could be considered dominant to the point of skewing the results. Further, the very stable and significant results of this study are clearly due to the combination of high interobserver reliability with a large number of observations ($N = 50{,}488$).

STUDY 2

Experimental Design

A second study was conducted in order to verify and extend the findings of the first study. The goal was to replicate the findings and to demonstrate an even stronger effect of the training program. Consequently, a modified experimental design was employed in order to eliminate the mitigating influence of the untrained control group.

The second study performed at Warehouse 2, used a within groups only comparison, that is, all 48 lift truck operators were trained at the same time and all received performance feedback. Comparisons were made only before and after training. In addition, the schedule of the second study was abbreviated because it was clear from the results of the first that less time and fewer observations would be sufficient to achieve stable and significant results. Each phase of the study (pre-training, post-training, and retention) therefore lasted 2 weeks.

Results

Figure 2 shows the mean error rates for the Pre-training, Post-training, and Retention phases. After training, there was an immediate 61% improvement in performance scores. Observations in the Retention phase, 3 months after the post-training observations, showed an additional reduction of 22% in mean error rates. This corresponded closely to the 24% additional gain found in Study 1.

The overall net improvement in mean error rates was 70%. This was stronger than

TABLE 5
PERCENT OF TOTAL OBSERVATIONS FOR INDIVIDUAL BEHAVIORS
(WAREHOUSE 1)

BEHAVIOR	NUMBER OF OBSERVATIONS	PERCENT OF TOTAL OBSERVATIONS
1. Warns trucks	392	<1%
2. Yields to trucks	617	1%
3. Warns co-workers	816	2%
4. Yields to co-workers	1,393	3%
5. Sounds horn at blind intersection	769	2%
6. Slows down at blind intersection	766	2%
7. Looks at blind intersection	731	1%
8. Looks in direction of travel	7,849	15%
9. Maintains moderate speed	6,383	13%
10. Avoids quick starts/changes of direction	5,141	10%
11. Keeps all body parts within truck	7,915	16%
12. Maintains forks in proper position	7,685	15%
13. Maintains balanced load	4,306	9%
14. Drives in reverse	5,725	11%
All behaviors	50,488	100%

that observed in Study 1 (44%) and was highly significant ($F [2, 16] = 307.75$, $p < .0001$). Post-hoc analyses using the Duncan test confirmed that mean errors at each phase of training program assessment were significantly different, all in the predicted direction. The very strong and stable effects were again attributed to the large number of observations ($N = 12{,}107$) and high interobserver reliability coupled with the modified experimental design which served to minimize experimental error.

Table 6 shows that large improvement occurred for all behaviors, including behaviors 11 and 14 which did not improve in Study 1. Overall pre-training error rates were lower

FIGURE 2
MEAN ERROR RATES FOR SUCCESSIVE PHASES OF THE TRAINING PROGRAM
(WAREHOUSE 2)

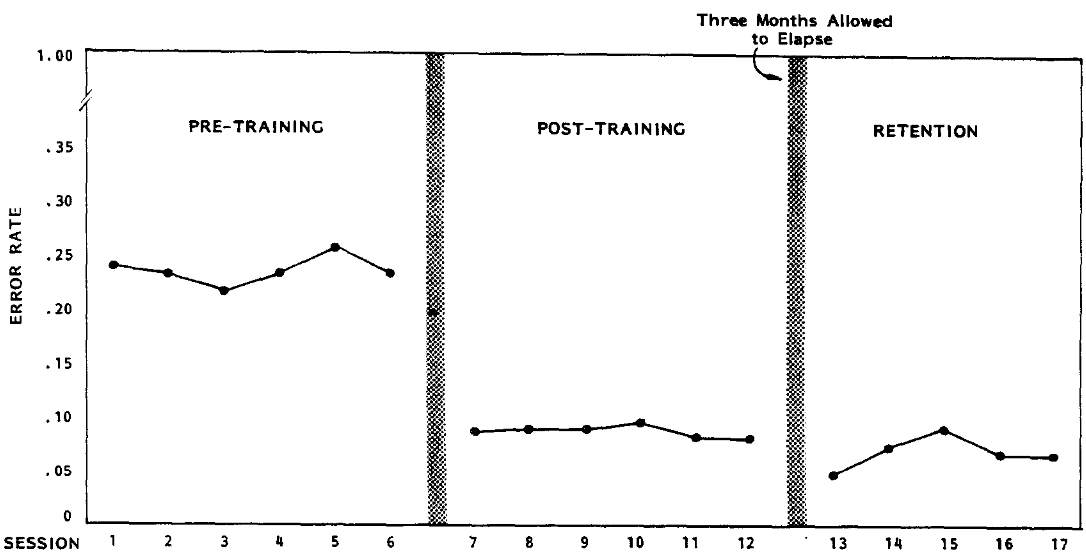

TABLE 6
PERCENT CHANGE IN ERROR RATES FOR INDIVIDUAL BEHAVIORS
(WAREHOUSE 2)

BEHAVIOR	PRE-TRAINING MEAN ERROR RATE	RETENTION MEAN ERROR RATE	PERCENT CHANGE IN ERROR RATE
1. Warns trucks	.18	.04	77%
2. Yields to trucks	.28	.11	61%
3. Warns co-workers	.64	.13	80%
4. Yields to co-workers	.64	.24	63%
5. Sounds horn at blind intersection	.31	.06	81%
6. Slows down at blind intersection	.63	.25	60%
7. Looks at blind intersection	.42	.13	69%
8. Looks in direction of travel	.30	.11	63%
9. Maintains moderate speed	.24	.09	63%
10. Avoids quick starts/ changes of direction	.14	.03	79%
11. Keeps all body parts within truck	.15	.02	87%
12. Maintains forks in proper position	.16	.07	56%
13. Maintains balanced load	.12	.04	67%
14. Drives in reverse	.17	.03	82%
All behaviors	.23	.07	70%

for the second warehouse compared to the first (.23 vs. .34) as were overall mean error rates during the final (retention) phase of both studies (.07 vs. .19). Improvement was generally stronger for all behaviors observed in Warehouse 2 despite the fact that operators there generally exhibited fewer errors prior to the introduction of training.

Table 7 shows that, similar to Study 1, no single behavior dominated the observed effects, with percent of total observations for each behavior ranging from about 1 to 15% with a median of about 9%. The distribution of total observations for Warehouse 2 is remarkably similar to that for Warehouse 1 (Table 5), which supports the impression that

TABLE 7
PERCENT OF TOTAL OBSERVATIONS FOR INDIVIDUAL BEHAVIORS
(WAREHOUSE 2)

BEHAVIOR	NUMBER OF OBSERVATIONS	PERCENT OF TOTAL OBSERVATIONS
1. Warns trucks	104	<1%
2. Yields to trucks	143	1%
3. Warns co-workers	121	1%
4. Yields to co-workers	196	2%
5. Sounds horn at blind intersection	490	4%
6. Slows down at blind intersection	488	4%
7. Looks at blind intersection	465	4%
8. Looks in direction of travel	1,819	15%
9. Maintains moderate speed	1,446	12%
10. Avoids quick starts/changes of direction	1,100	9%
11. Keeps all body parts within truck	1,834	15%
12. Maintains forks in proper position	1,799	15%
13. Maintains balanced load	1,040	9%
14. Drives in reverse	1,062	9%
All behaviors	12,107	100%

the functions and activities at the two warehouses were quite comparable.

DISCUSSION AND CONCLUSION

These two studies made use of a behavioral (work) sampling procedure to obtain objective data about work practices that correlate with injury risk. Although the technique has been rarely used in the past, it proved invaluable for objectively measuring the effects of safety training. It also proved useful in serving as the basis for training program development and providing performance feedback.

The results of the two studies show that a well designed and administered occupational safety training program, emphasizing safe work practices and derived from a true assessment of need, can be effective in improving on-the-job behavior. Even better performance can be achieved by following the training with a program based on goal-setting and performance feedback supplemented with informal peer group modeling. This is similar to the conclusions reached by Locke (1980) in an assessment of past research. Further, the improved performance can endure well beyond the cessation of daily performance monitoring and feedback. The explanation for the enduring effects of the program appears to be that habits were changed due to continued practice in safe work procedures, coupled with a redefinition of group norms sustained through peer modeling and continued management support.

These conclusions support those of Zohar (1980) who suggested that informal influences, such as peer modeling and management support, are the ultimate and most practical types of reinforcement for *sustaining* modified safe work practices following cessation of formal training programs.

REFERENCES

Cohen, A., Smith, M. J., & Anger, W. K. (1979). Self-protective measures against workplace hazards. *Journal of Safety Research, 11*, 121–131.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37–46.

Das, B. (1982). Effects of production feedback and standards on worker productivity in a repetitive production task. *IIE Transactions, 14*, 27–37.

Goldstein, I. L. (1974). *Training: Program development and evaluation*. Belmont, California: Brooks/Cole.

Komaki, J., 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., 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.

Locke, E. A. (1963). Toward a theory of task motivation. *Organizational Behavior and Human Performance, 3*, 157–189.

Locke, E. A. (1980). Latham versus Komaki: A tale of two paradigms. *Journal of Applied Psychology, 65*, 16–23.

McCormick, E. J., & Sanders, M. S. (1982). *Human factors in engineering and design* (5th ed.). New York: McGraw-Hill.

Meister, D. (1976). *Behavioral foundations of systems development*. New York: John Wiley.

Rockwell, T. H. (1959). Safety performance measurement. *The Journal of Industrial Engineering, 10*, 12–16.

Smith, M. S., Anger, W. K., & Uslan, S. S. (1978). Behavior modification applied to occupational safety. *Journal of Safety Research, 10*, 87–88.

Surry, J. (1969). *Industrial accident research*. Toronto: Labour Safety Council, Ontario Department of Labour.

Tarrants, W. E. (1980). *The measurement of safety performance*. New York: Garland.

Zohar, D. (1980). Promoting the use of personal protective equipment by behavior modification techniques. *Journal of Safety Research, 12*, 78–85.

Zohar, D., Cohen, A., & Azar, N. (1980). Promoting increased use of ear protectors in noise through information feedback. *Human Factors, 22*, 69–79.