Construction Management and Economics (1994) **12**, 67–78

Improving safety by the modification of behaviour

A.R. DUFF, I.T. ROBERTSON, R.A. PHILLIPS and M.D. COOPER

Department of Building Engineering and *Manchester School of Management, University of Manchester Institute of Science and Technology, Manchester, UK*

Received 4 January 1993 and accepted 27 April 1993

Safety improvement in the construction industry will only be achieved if all concerned in the operation of construction sites change their behaviour. This article describes the development and effects of behaviourally-based management techniques in improving site safety. Goal-setting and feedback methods were developed and tested on six construction sites in the Northwest of England. A longitudinal research design was adopted, in which measures of safety performance were taken before, during and after the application of these methods. The measures included four categories of measurement: access to heights; site housekeeping (tidiness); scaffolding; and use of personal protective equipment. Three of these categories were used for experiments with a variety of goal-setting and feedback interventions, while the fourth was used as a control. The results show that: safety behaviour can be objectively and reliably measured; goal-setting and feedback can produce large improvements in safety performance; commitment of site management appears to enhance their effectiveness. The work was financed by the Health and Safety Executive, and a further contract to develop the techniques into a basic part of safety management for several large construction companies has begun.

Keywords: Safety management, measurement, goal-setting.

Introduction

In 1976 the UK Chief Inspector of Factories made a then controversial forecast, 'Unless there is a radical change in the effectiveness of accident prevention in the (construction) industry, about 2000 men will die and about 400 000 will be injured over the next 10 years' (Dalton, 1989). The construction industry has an unenviable safety record. On past evidence it is likely that about 150 people will die by accident each year. A further 2500 to 3000 will be seriously injured. In addition 30 000 to 40 000 will suffer lost time injuries of more than 3 days (Davies and Tomasin, 1990).

A review of previous research reveals that attempts to improve this situation have not been successful. These attempts have included 'blitz' inspections by the Health and Safety Executive. During 1987–88 inspections of over 2000 sites were conducted. These inspections revealed a worrying picture with one third of site agents and supervisors having inadequate knowledge of basic health and safety requirements. Most importantly during the period of the campaign there was no decrease in the number of deaths or serious injuries (Health and Safety Executive, 1988). Other kinds of interventions, such as posters and informational safety campaigns, designed to improve safety have not been consistently successful (e.g. Kay, 1978; Arscott, 1980; Wilson, 1989). This article describes the development and implementation, and reports the results, of a set of behaviourally-based management procedures designed to improve safety on construction sites.

Behavioural approaches

Previous research (e.g. McAfee and Winn, 1989) has shown improvements in safety behaviour in several industries, through the use of psychologically-based techniques. These techniques, generically termed 'applied behaviour analysis', involve the use of goal-setting and performance feedback. It is also common for extrinsic reward (reinforcement) to be incorporated, but since reinforcement was not employed in the research described in this article this aspect of applied behaviour analysis is not of direct interest.

The literature on goal-setting as a procedure for managing behaviour is voluminous. Wood, Mento and Locke (1987), for example, reviewed nearly 200

0144–6193 © 1994 E. & F.N. Spon

Copyright © 2001. All Rights Reseved.

scientific, empirical studies. Goal-setting theory hypothesises that goals are the immediate, though not sole, regulators of human action and that performance will improve when goals are hard, specific and accepted by the actor. Thus there is expected to be a positive correlation between goal difficulty and performance. In general the literature on goal-setting supports these propositions and provides clear guidance on how to operationalize the theory to good effect (Locke and Latham, 1990).

The research literature on the role of feedback in determining performance effectiveness is clear in indicating the positive effect of knowledge of the results of one's behaviour. Reviews of the research on feedback (Guzzo, Jette and Katzell, 1985; Kopelman, 1986; Algera, 1990) demonstrate that performance is enhanced when management techniques that provide clear feedback of performance-related information are used.

There have been several attempts to improve the poor accident record by raising operatives' safety consciousness through the use of safety poster campaigns etc. Such campaigns are generally ineffective, as illustrated by Saarela *et al.* (1989), and do not make a lasting impact on the accident/injury rate. Previous research (Shimmin *et al.*, 1981) has shown that in a sample of accident victims two-thirds considered their accidents to be avoidable, suggesting that, in the view of the workers themselves, something could be done to reduce accidents. Accident causes identified by 'victims' included a high proportion of references to inappropriate behaviour or equipment usage. Another frequently cited factor was that the reward system rewarded speed of work but did not reward safe working procedures. This suggests that restructuring the information feedback and goal-setting systems may influence the specific behaviours involved in causing accidents which is exactly the approach at the core of this research.

McAfee and Winn (1989) showed that safety behaviour can be improved by systematically monitoring safety-related behaviour and providing feedback in conjunction with goal-setting and/or training. In combination, goal-setting and performance feedback provide a powerful management tool by focusing on job tasks. Goal-setting affects performance by: directing the attention and actions of the individual/group; mobilizing effort; and increasing motivation. Setting difficult, yet achievable goals, and providing performance feedback in relation to them, can influence behaviour, if employees are committed to the goals.

When the research described in this paper began, these techniques for changing behaviour, developed from the specific approaches to human motivation mentioned above, had already been shown to be of value in safety (e.g. Zohar and Fussfeld, 1981) and productivity (O'Brien *et al.*, 1982). Some work in construction-related industries and in other countries had also been done (e.g. Komaki, 1977; Rhoton 1980 and Chokkar and Wallin, 1984). No previous attempt had been made to apply the techniques in the UK construction industry.

Behaviour change, of the kind needed to improve safety, is sometimes the target of training programmes. As already noted, personnel on construction sites in the UK are not always familiar with safety requirements. It is apparent that although training alone may not bring about improvements in safety it may be of some value. As far as safety training is concerned research findings suggest that training material specifically linking the hazards of the situation with harm that might be caused is effective (Hale, 1984).

Research goals

The general aim of the research reported in this paper was to evaluate the benefits of using goal-setting, feedback and training techniques to improve safety on construction sites in the UK. In order to test the effectiveness of these techniques, an accurate and reliable measure of safety performance was required which could be used before, during and after their application. The specific objectives of the research, therefore, were to:

1. develop and test a method of measuring safety performance on construction sites;
2. use the method to evaluate specific management procedures, based upon proven techniques for changing work behaviour, to improve site safety.

Research method

A longitudinal research design was adopted and differences in safety levels before and after the use of the techniques were evaluated. In addition an experimental approach was taken so that any differences in safety performance could be attributed to the intervention techniques rather than extraneous factors. The steps in the research are described below.

Study design

An experimental design, with withdrawal, was utilized on six construction sites in the Northwest of England. The use of this design is linked to the need for the research to provide clear, interpretable results concerning the effectiveness of the interventions.

Withdrawal involves alternating the application of the intervention (e.g. introducing goals and a feedback chart) and removal of the intervention (the withdrawal

Copyright © 2001. All Rights Reseved.

phase). In such a design the target behaviours are measured before the intervention programme begins to provide baseline data. The next phase involves introducing the intervention and taking measures to test its effect. This is followed by the withdrawal phase in which conditions revert to 'normal' and attention is focused on whether any behaviour change (e.g. improved safety performance) is retained or not.

The experimental aspect of the design involved collecting data on three experimental categories of behaviour, related to scaffolding, housekeeping and access to heights, and a control category related to personal protective equipment (PPE).

Safety performance measurement

Developing an objective and quantifiable method of safety measurement was achieved by identifying contributory factors in the chain of events which cause accidents. In order to produce a comprehensive list of unsafe items (situations or behaviours), a detailed literature review of construction journals, HSE publications, construction safety manuals and accident records was undertaken. From an analysis of various types of fatal and major injuries reported to the Health and Safety Executive (1988), 99 of these items were selected and incorporated into a questionnaire, which formed the basis for a survey (n = 194) of construction personnel to determine the perceived importance, or risk level, of each item in accident causation. Respondents rated each item on three scales: frequency of occurrence, likelihood of a resulting accident and probable severity of such an accident, and these dimensions were combined to determine the relative importance of these items. A total of 24 items were selected and incorporated into a safety audit checklist, used to evaluate the safety performance of construction sites in the experiments to change safety behaviour. Items were categorized to form four composite measures of safety: scaffolding; access to heights; housekeeping and personal protective equipment (see Appendix 1).

Previous research involving measures of safety performance has normally utilized an 'all or nothing' (AON), i.e. either 100% safe or 100% unsafe, measure of safety performance (Komaki *et al.*, 1978). However, a 'proportional rating scale' (PRS) (Cooper *et al.*, 1991) is better suited to cope with the construction industry environment. For example, a scaffold with only 75% of the required toe-boards correctly fixed would be assessed as 25% unsafe on that particular measurement item, rather than just 'unsafe'. This permits changes in safety performance to be measured with increased sensitivity.

In order to reflect the sensitivity of this measure, an 11-point rating scale was employed A 'Not seen' result was used to record occasions when there was no evidence of the item, in either safe or unsafe condition. The scale is anchored using expressions of amount which have been validated by previous research undertaken by Bass *et al.* (1974) and records the proportion of unsafe situations or behaviours. It records unsafe situations/behaviours as a proportion of total opportunities to be safe. Further details of the advantages of this proportional scale, rather than an AON scale are given in Cooper *et al.* (1991).

Intervention procedures

Three different kinds of intervention (goal-setting, feedback and training) were developed for the research.

1. *Goal-setting.* A procedure was developed to introduce goal-setting as a tool to influence safety performance on construction sites. Goal-setting scripts were devised which included an explanation of who the experimenters were, how the safety performance levels were determined, what the current levels of safety performance of the particular intervention category were, and how frequently performance feedback would be given.
2. *Feedback charts.* Performance feedback charts, capable of showing graphically 42 weeks of measures, were designed and mounted in locations visible to all site personnel. The vertical axis indicated safety performance, while the horizontal axis indicated the week number. Three coloured tapes were used to plot safety performance levels on each chart. In addition, three coloured tapes with 'Target Level' printed upon them were used to indicate the goal to be reached.
3. *Training.* The principal objective of the training intervention was to ensure that the site personnel understood fully the basis of the goal-setting and feedback activities and, in particular, the items which were contained in the measurement 'check list'. A standard training package was devised for each experimental category. Each training package consisted of a 15-min slide/tape presentation with slides depicting actual unsafe situations and behaviours found on construction sites, and a commentary that emphasized the correct manner in which to behave, followed by a summary slide re-emphasising the 'do's and don'ts'. The slides focused, in turn, on each item in the category, specifically linking the hazards of a situation to the harm that might be caused. Some slides were purpose made and others were selected from safety training materials from the Construction Industry Training Board.

Meetings were held in site canteens with groups of

Copyright © 2001. All Rights Reseved.

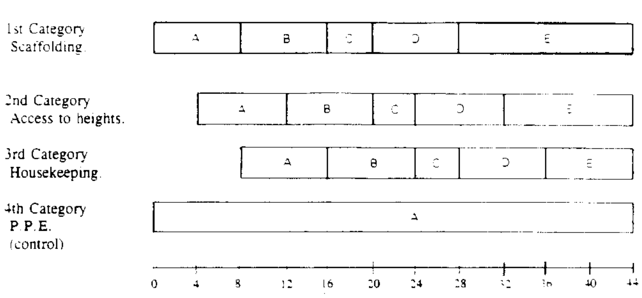
approximately 30 or 40 operatives at a time. On some large sites more than one meeting was necessary to ensure the attendance of all operatives. Upon completion of the presentation, questions were invited from the audience. Points of interest from the training were fully discussed until consensus was achieved as to the correct manner in which to behave. Commonly, questions were raised regarding the provision of resources to allow and support safe systems of working, and the site management's role.

In some instances, due to work-related activities (e.g. concrete pouring), not all of the operatives were able to attend the site meetings. To overcome these difficulties, training display boards were devised that presented the training material in pictorial form, and were placed next to the safety performance feedback chart.

The entire presentation, including introduction by a researcher, the slide/tape showing, questions from the audience, and when appropriate the goal-setting process, took no more than 30 min. This was essential in order to minimize disruption to work activities.

Data collection

Data collection not used in the experiments was carried out for an initial period of up to 12 weeks to dissipate any effects caused by the presence of observers on site. After this, the procedure involved: measuring safety performance for 8 weeks to provide a base-line measure; introducing the intervention for 8 weeks and measuring the change in performance; terminating the intervention (withdrawal) and measuring performance for a further 4 weeks. The intervention and withdrawal cycle was then repeated (see Fig. 1). It is important to note that interventions were never aimed at on the control category (PPE). The expectation was that, if the interventions were effective, safety performance would



Key:
A Baseline period.
B 1st Intervention period.
C 1st Withdrawal period.
D 2nd Intervention period.
E 2nd Withdrawal period.
PPE Personnel protective equipment

Figure 1 Experimental design

improve in the experimental categories during the various stages of the study, but remain constant in the control category.

As Fig. 1 shows, this entire intervention/measurement process was carried out for three categories of activity: scaffolding, access to heights, and general housekeeping, while the fourth category (PPE) was monitored as a control.

The final research design investigated combinations of goal-setting, training and feedback on six different sites (see Fig. 2).

	Site	Feedback	Training	Goal-setting
1	Plastic processing factory	✓	X	✓
2	New hospital wards	✓	X	✓
3	Large office & retail development	✓	✓	✓
4	City centre apartments	✓	✓	✓
5	Offices in city centre	✓	✓	X
6	Bank premises in city centre	✓	✓	X

Figure 2 Sites and intervention protocols

Independent observers (i.e. observers without knowledge of the specific aims of the research) were employed to conduct the safety audits, in order to reduce the likelihood of experimenter bias from expectations born of familiarity with the experimental design. The observers were trained for a period of 4 weeks in the use of the safety performance measure.

Observations on each site were taken three times a week at differing times of the day and days of week, to overcome any systematic time of day or week effects and to avoid operative and management expectations affecting the results. Observation sessions took approximately 45 min on each site, and were never taken twice on the same day. Observations on each site were all undertaken by one observer, as each observer was allocated to specific sites throughout the duration of the research. Consensual drift (i.e. where observers appear to drift from the original consensus about the items, systematically and together, across observations) and inter-observer reliability were monitored regularly by random checking of observations by a researcher who would independently complete a measure for the same site. Using the percentage agreement method (Komaki, 1978) these checks showed 92% agreement between observers and researchers.

In the very few instances, due for example to ill health, that the independent observers were unable to carry out their observations, the researchers conducted them. Where possible these were carried out by two researchers jointly, to minimize bias.

Copyright © 2001. All Rights Reseved.

Quantitative results and analysis

The procedures adopted in this study produced safety performance data for six construction sites over a period of 40 weeks. The analyses of these data, presented below, focus on two key questions:

1. are there statistically significant changes in safety performance in the experimental categories (scaffolding, housekeeping, access to heights), coupled with no changes in the control category (PPE)?
2. does safety performance decrease when goals and/or feedback are withdrawn, and then improve again when the interventions are reintroduced?

Taken together these questions provide a firm basis for assessing the impact of the interventions used in the study.

An answer to the first question posed above can be seen most directly by examining safety levels after the intervention, compared with the baseline. With the design for this study involving six sites, two separate intervention phases, four safety categories and three types of intervention, a large number of pre-post differences can be examined. Each separate pre-post difference will give only a partial view of the overall effects caused by the interventions in the experiment and it was necessary to utilize an analytical procedure which could cumulate the results and provide an overall view of the data. The analytical approach adopted meta-analysis procedures (Hunter and Schmidt, 1990) as an appropriate way of quantitatively cumulating the results obtained from the intervention protocols (i.e. combinations of goal-setting, feedback and training) used in the research.

Meta-analysis requires that a common statistic (which indicates the size of the effect being examined) is identified in order to facilitate the accumulation of results across or within studies. This may be achieved by using the results of the studies to calculate a statistical effect-size indicator such as 'd' (Cohen, 1987), the standardized difference between group means, divided by the within group standard deviation.

An an example, in the case of the current study it is possible to calculate a 'd' statistic to indicate the improvement, or otherwise, between baseline safety performance in each category and performance after the first intervention. This is done by calculating the mean safety performance score for the last 4 weeks of the baseline period (the pre-score) and the mean safety performance score for the final 4 weeks of the intervention (the post score). The relevant 'd' statistic is that given by substituting into the formula:

$$d = (\text{pre score} - \text{post score})/\text{sd(pooled)};$$

where sd(pooled) = the standard deviation of the full set of scores.

To analyse the results for this study d-statistics were calculated for two different phases of the research across all sites and categories. The first set of d's (phase one) were derived by comparing the means for the final 4 weeks of the baseline with the final 4 weeks of the first intervention.

The second set of d's (phase two) were calculated by comparing the means for the 4-week withdrawal period with the final 4 weeks of the second intervention. The d's from phase one of the study focus on the effect of the first intervention period and relate to the question of whether or not the intervention had an effect. The d's from the second phase of the study focus on the effect of re-introducing the intervention after a period of withdrawal (i.e. after the withdrawal of feedback).

Table 1 provides meta-analysis of the results by cumulating all of the results obtained for phase one and phase two across all sites for the whole study period, for all of the intervention categories combined. Results for the control category (PPE) are also given. These aggregate results do not show separate effects for the first intervention against the original base line, nor do they show separate, subsequent effects measured against performance during the withdrawal period (which may be thought of as a second baseline). In this general table of results both types of effect are combined to provide a single overview of the impact of the interventions, during the course of the whole study.

Table 1 Global meta-analysis results for experimental and control categories compared

	Number of observations	Number of effect sizes	Mean observed effect size 'd'	Mean corrected effect size 'd'	Variance corrected for sampling error	Boundary of credibility interval (alpha = 90%)
Experimental categories (scaffolding, access to heights, housekeeping)	771	35	0.4452	0.4594	2.6976	−1.643
Control category (personal protective equipment)	267	12	0.1092	0.1109	0.2036	−0.467

Copyright © 2001. All Rights Reseved.

Table 1, and subsequent tables, provide the same types of information:

1. 'Number of observations' refers to the number of separate observations made by the observers (using the observer check list) that have been included in the analysis presented.
2. 'Number of effect size' refers to the number of separate d values included in the analysis. For example in Table 3 the results for the control category (PPE) are based on 12 separate d's (i.e. PPE on six sites for two different intervention phases ($6 \times 2 = 12$).
3. 'Mean observed effect size' gives the average effect size for the d values in the relevant row of the table.
4. 'Mean corrected effect size' gives the mean effect size corrected for unreliability of measurement.
5. 'Variance corrected for sampling error' gives the variance in d values (for the relevant row) after correction for variation caused by sampling error.
6. 'Boundary of credibility interval' gives the boundary of the 90% credibility range for d. This is calculated using a procedure similar to the normal procedures for fitting a confidence interval around a mean score (see Whitener, 1990).

Table 2 shows these results subdivided in various ways to provide clear indications of separate effects for various aspects of the experiment. The top section of Table 2 gives the results separately for each study phase. The second part of the table gives the results further subdivided according to categories of safety performance investigated.

The boundary of the credibility interval provides a signal as to whether or not the mean corrected d gives a generalizable estimate of the population effect size. If the boundary of the credibility interval does not take in zero, the mean d is probably an indication of the effect that would be obtained in other situations similar to the ones examined in the research (see Whitener, 1990). (Technically, when the boundary of the credibility interval includes zero *and* the variance after correction for sampling error is small, the d's involved in the calculation were probably drawn from a single, homogeneous population, see Whitener, 1990; Hunter and Schmidt, 1990.)

The results in Table 1 show a mean corrected d for the intervention (to two decimal places) of 0.46. Anything around 0.5 or above is a large effect. For the control category a small d of 0.11 was found. These initial, very positive results are complicated by the fact that the boundary credibility value for the global d-statistic given in Table 1 includes zero. This suggests that the effect is not generalizable and that, for some of the experimental settings, different sizes of effect can be expected. In other words, the individual effect sizes are not from a homogeneous sample, a result which is unsurprising in view of the different phases, categories and intervention protocols.

The further analysis given in Table 2 provides some strong clues about why the variations in effect sizes occur. It is clear from Table 2 that the first intervention has a very strong positive effect (phase one mean $d = 1.53$), yet during the second phase of the project, safety levels tend to fall back from the improved levels

Table 2 Results broken down by study phase and category

	Number of observations	Number of effect sizes	Mean observed effect size 'd'	Mean corrected effect size 'd'	Variance corrected for sampling error	Boundary of credibility interval (alpha = 90%)
(i) **Phase period**						
Phase One	398	18	1.4857	1.5335	2.0627	−0.305
Phase Two	373	17	−0.6652	−0.6864	0.7671	−1.808
(ii) **Category × phase**						
Scaffolding						
Phase One	136	6	1.4862	1.5034	2.3954	−0.478
Phase Two	131	6	−0.6652	−0.6729	1.6329	−2.309
Access to heights						
Phase One	129	6	2.0818	2.1881	2.5225	0.155
Phase Two	130	6	−0.8725	−0.9171	0.3009	−1.619
Housekeeping						
Phase One	133	6	0.9072	0.9382	0.5233	0.012
Phase Two	112	5	−0.4245	−0.4389	0.1132	−0.870

Copyright © 2001. All Rights Reseved.

reached in phase one (phase two mean d = −0.67). This result, that the phases of the study show different effects, is supported by the noticeable decrease in the variance of effects when the results are grouped in this way. The aggregate variance in Table 2 (for intervention categories) is 2.6698, whereas the average variance for the two phases shown separately in Table 2 shrinks to 1.4149. On the other hand the lower boundaries of the credibility intervals still include zero, suggesting that further factors may be moderating the effect of the interventions.

Table 2 shows a very consistent picture in relation to phase one effects, which are large and positive for all categories. In two of the three-intervention categories the boundary of the credibility interval exceeds zero (access to heights and housekeeping). These results suggest that the first phase of the intervention has had a strong positive effect, although there is considerable variation across sites. By contrast the second phase of the intervention programme does not produce further positive effects. The trend is for safety levels to decrease but this effect is not strong enough, nor consistent enough, to support the conclusion that there is a decrease during phase two. There is, however, certainly no evidence of an overall increase during phase two.

Before attempting to explain these results, it is useful to reflect upon the procedures of goal-setting and feedback and the ways in which they might be expected to affect performance. When goals are set they focus attention upon certain areas of safety performance, represented by the items included in the safety measures. It is expected that this attention results in a direction of effort towards improving behaviour. The result of the effort will be a progressive improvement in the safety condition of the site which will register in the safety measures being recorded. In the study described here, at the beginning of the week after the measures were taken, the feedback charts were updated, communicating improvement to site personnel and reinforcing their efforts to improve further towards their goals. In any given setting these efforts may be applied directly, in the case of operatives, or indirectly, in the case of managers and supervisors.

It can be seen that there is normally a time lag between the initial procedure of goal-setting, with its subsequent feedback, and the achievement of substantial improvement. The effects may be cumulative and self-reinforcing, developing an internal momentum which only dissipates when further improvement is impossible or not valued.

Withdrawal of the feedback removes the reinforcement and reward for good performance and, ultimately, eliminates the special attention to the areas of performance upon which goal-setting and feedback is focused. Behaviour in relation to these areas deteriorates, as does the resulting safety condition of the site. Once more it can be seen that there is a time lag before the effect is evident in the safety measures.

Due to these time lags, the results of the intervention-withdrawal-intervention-withdrawal cycle used in this research programme might be expected to result in a pattern of performance similar to the graph in Fig. 3.

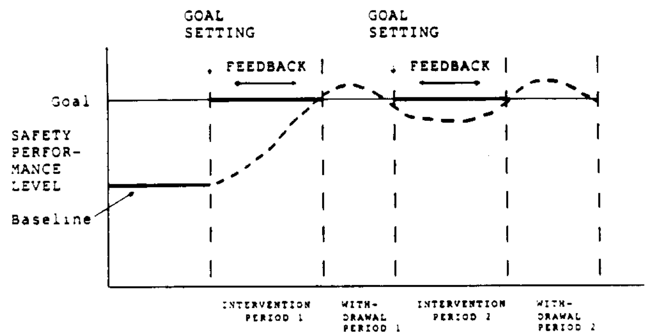


Figure 3 Expected effect of two-cycle intervention pattern

The results obtained did, in fact, follow this pattern.

Graphical results for this repeated pattern of intervention and withdrawal are typified by Fig. 4. The general pattern, based upon a visual appraisal of all of the graphical results for the study, may be summarized as follows:

1. most measures (14 out of 18) showed improvement by the second half of the first intervention period, when compared with the performance during the last 4 weeks of the baseline period;
2. performance during the first withdrawal generally continued to rise with nine of the 18 results showing an improvement, compared with the second 4 weeks of the intervention period, four results remained steady, and five cases showed only small falls;
3. performance during the second intervention period, measured during the second 4 weeks, generally showed a decline with only two out of the 18 cases showing a slight improvement and 11 showing a decline when compared with the first withdrawal;
4. the first 4 weeks of the second withdrawal period showed little change, with five cases showing virtually no change at all and the remainder small falls or rises, when compared with the second 4 weeks of the second intervention.

It can be seen, therefore, that the overall pattern supports quite well the hypothesis suggested by Fig. 3.

These results could also suggest another possibility: namely that the disappointing result of the second intervention phase, when compared with the very positive effect of the first, could result from disaffection

Copyright © 2001. All Rights Reseved.

of the operatives with an activity which may have, by then, appeared to them to be clearly experimental. To put it in the terms that they might have applied, 'If it was working (as it clearly was at the beginning), why mess about with withdrawal phases?' This possibility was not recognized in time to test its validity.

The results are further disaggregated in Table 3 to show the relative effects of the different intervention protocols referred to in Fig. 2. Further statistical analysis of these showed that goal-setting and feedback seemed to have the strongest influence, with training, goal-setting and feedback slightly weaker and training and feedback the weakest. These very tentative conclusions derive from data aggregated across all sites, categories and phases. However, the differences are neither clear enough, nor significant enough in statistical terms, to be able to report this result with confidence.

Some qualitative results

A number of other interesting and potentially important results arise from an analysis of the wide range of qualitative and anecdotal evidence collected during the project.

Management commitment

Although no attempt was made, as an integral part of the research programme, to evaluate the effect of management commitment, it became apparent that this was

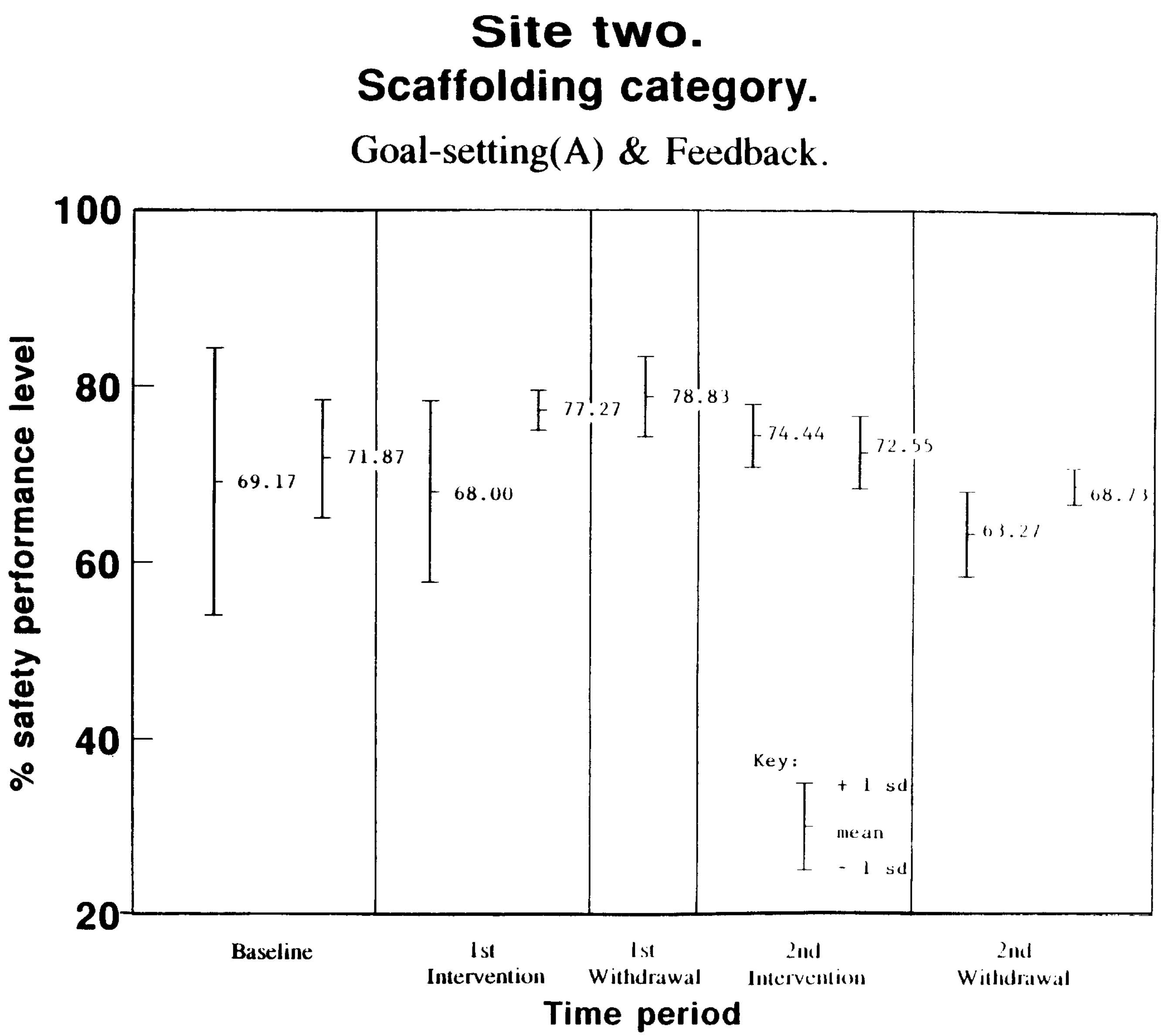


Figure 4. Site two: scaffolding category

Copyright © 2001. All Rights Reseved.

Table 3 Results broken down by category, protocol and phase

	Number of observations	Number of effect sizes	Mean observed effect size 'd'	Mean corrected effect size 'd'	Variance corrected for sampling error	Boundary of credibility interval (alpha = 90%)
Category × protocol × phase						
Scaffolding:						
Tg, Gs & Fdbk (Phase One)	40	2	0.8656	0.8757	1.5035	−0.694
Tg, Gs & Fdbk (Phase Two)	40	2	−1.3336	−1.3490	0.2507	−1.990
Gs & Fdbk (Phase One)	49	2	3.1620	3.1986	4.5479	0.469
Gs & Fdbk (Phase Two)	46	2	−1.2798	−1.2946	2.1908	−3.189
Tg & Fdbk (Phase One)	40	2	0.6163	0.6235	0.2370	0.004
Tg & Fdbk (Phase Two)	21	1	−1.1113	−1.1242	*	−1.124
Access to heights:						
Tg, Gs & Fdbk (Phase One)	44	2	1.6854	1.7715	*	1.772
Tg, Gs & Fdbk (Phase Two)	46	2	−1.7839	−1.8749	*	−1.875
Gs & Fdbk (Phase One)	41	2	2.4912	2.6184	0.1415	2.137
Gs & Fdbk (Phase Two)	40	2	−1.1199	−1.1771	*	−1.177
Tg & Fdbk (Phase One)	42	2	3.0197	3.1738	2.5020	1.149
Tg & Fdbk (Phase Two)	24	1	0.0441	0.0464	*	0.046
Housekeeping						
Tg, Gs & Fdbk (Phase One)	47	2	0.3104	0.3210	0.6030	−0.673
Tg, Gs & Fdbk (Phase Two)	47	2	0.0719	0.0744	0.7032	−0.999
Gs & Fdbk (Phase One)	44	2	1.3908	1.4382	*	1.438
Gs & Fdbk (Phase Two)	45	2	−0.1568	−0.1622	1.5910	−1.777
Tg & Fdbk (Phase One)	47	2	1.5190	1.5709	*	1.571
Tg & Fdbk (Phase Two)	23	1	1.8415	1.9043	*	1.904
Control category (PPE)						
Tg, Gs & Fdbk (Phase One)	40	2	−0.0054	−0.0054	*	−0.005
Tg, Gs & Fdbk (Phase Two)	40	2	0.7082	0.7190	*	0.719
Gs & Fdbk (Phase One)	49	2	−0.0522	−0.0530	*	−0.053
Gs & Fdbk (Phase Two)	46	2	0.0752	0.0763	0.0037	−1.309
Tg & Fdbk (Phase One)	40	2	0.3170	0.3767	0.6047	−0.619
Tg & Fdbk (Phase Two)	33	2	−1.4516	−1.4736	*	−1.474

Tg = training; Gs = goal setting; Fdbk = feedback.
* The formulae used to estimate variance due to sampling error may sometimes produce estimates for sampling error that are larger than the original observed variance. This occurs most often when the number of effect sizes are small. When this has happened * is given for the corrected variance; note that although such variances cannot be reliably determined it is likely that they are very close to zero.

probably having a significant impact upon some of the results.

It was evident on some sites that the site management, despite support for the project from senior company management, were not fully committed to the research activity. For example, in spite of oral assurances that they would do so, there were two sites upon which the site management did not attend any of the training or goal-setting sessions. This could have been interpreted by the operatives as a lack of commitment by the management to the objectives of the research. There is little doubt that the day-to-day interaction between management and operatives on these sites also suggested a lack of managerial commitment to the research. Additionally, there were occasions when management demonstrated overt hostility to the research in disputing the validity of the safety measurement. It is relevant to note that the two best performing sites overall were those where management attended all the meetings with operatives at the commencement of both first and second interventions.

While the evidence is not conclusive, it does suggest that management commitment is a factor which should be objectively measured, and its effect evaluated, during any extension of this research.

Operative commitment

Operative commitment to the general aims and activities of the project was a factor which caused some concern

Copyright © 2001. All Rights Reseved.

prior to the commencement of the intervention programme. It was anticipated that a lack of commitment could, for example, result in rejection of the feedback and, in particular, vandalism of the feedback charts. The fact that this did not happen on *any* of the sites may indicate operative enthusiasm for the feedback of such safety performance data. This was also seen in the interest of operatives when feedback was posted on the charts. Further evidence came from expressions of operative support on an occasion when there was considerable disagreement between management and researchers about the validity of a particularly poor set of safety performance results. It was clear, from their comments, that the operatives were far more inclined to accept the research assessment of the safety of their site, than were the site management.

The general impression of the research team was that a sense of ownership of the project results, and in particular the feedback charts, developed in the site operatives as the project progressed. This impression was strengthened by quantitative data obtained through a questionnaire study (Bathgate, 1991) which revealed that operatives on one site wished to see such safety performance charts on all construction sites.

Conclusions

The research conducted in this project has shown that:

safety behaviour can be objectively and reliably measured, without excessive use of managerial or supervisory resource, producing performance data which could be used in many different safety management strategies;

goal-setting and feedback can be used to produce large improvements in safety performance, at least in the short term; the question of whether the improvement can be maintained is still to be resolved (see section on Further Work);

re-intervention (training and/or goal-setting) after withdrawal of the feedback, did not produce the same level of improvements as the first interventions, perhaps because there is a time-lag effect in the causal relationship between intervention and improvement, due to the inertia inherent in the personnel/site system;

commitment of site management appears to enhance the effectiveness of the goal-setting and feedback approach.

Three different combinations of training, goal-setting and feedbacks were tested. There was no conclusive difference in the results of the three intervention protocols, but they did suggest that goal-setting and feedback was better than feedback alone, and that the addition of training did not offer any benefit.

A full report of this research will be published by HSE and will be available through Her Majesty's Stationery Office during 1993, entitled: *Improving Safety on Construction Sites by Changing Personnel Behaviour*.

It is clear that the methods of safety improvement used during this research have had a positive impact and it is pertinent to contrast this with the minimal effects that previous research has recorded for other interventions such as information safety campaigns (Saarela, 1989) and safety training (Hale, 1984).

Further work

The interventions in this study were conducted by external agents (i.e. members of the research team) not employed by the construction companies concerned. Managerial commitment would be more evident if the training, goal-setting and feedback were conducted by company personnel. This could very probably result in even more powerful effects than those achieved in this research. Further work, also supported by HSE and designed to test this, is just beginning, with the enthusiastic commitment of all the contractors involved in this research. It has the following objectives:

1. develop the implementation methods of these techniques from a research based programme to a practical and effective management tool;
2. measure the relationships between management commitment, operative commitment and safety performance;
3. investigate the possible influences of organizational safety climate, management style, organizational structure (etc) on safety performance;
4. measure the cumulative effect of continuous and consistent efforts to improve safety through goal-setting and feedback.

Postscript

Feedback may be seen as a form of reinforcement, in which good performance is rewarded and poor performance punished, through its impact on such factors as self-esteem and job satisfaction. The impact is heightened by making public the results of such performance, by displaying them where they may be seen. There was no attempt, in this research project, to target individuals, and hence to reward individuals or to apportion blame for sub-standard performance. The approach was to consider the site and its personnel (operatives and management) as a unit, and to give the feedback to all personnel. The philosophy is that it is the responsibility of everyone to ensure safe working practices and.

Copyright © 2001. All Rights Reseved.

conditions. The theme of the message was 'Almost anyone can do something about it, even if only to remind those directly responsible'. This message operated in a variety of ways, ranging from supervisor instruction to improve an unsatisfactory situation, and therefore safety score, to peer pressure applied through teasing of colleagues about an unsatisfactory safety level shown in the feedback.

A clear indication of the way in which this approach works came in an incident on one site. Because of deficiencies in the scaffolding, which resulted in part from an inadequate specification and the refusal of the main contractor to pay for anything better, the safety measure for the scaffolding category was repeatedly sub-standard. After a time, and several disagreements between the main and scaffolding contractors, the director of the scaffolding company instructed that the situation be remedied. This was carried out at the sub-contractor's own expense, because 'I'm not having my company's name associated with a performance measure indicating sub-standard scaffolding, even if the main contractor won't pay to have the job done right'.

Acknowledgements

The authors wish to acknowledge their indebtedness to those individuals and companies who have made this project possible, including the two observers, John Garnett and Jeremy Hollingworth, for their conscientious approach towards the data collection, and the following construction companies for offering us unfettered access to their sites:

Alfred McAlpine (NW) Construction Ltd
Balfour Beatty Building Ltd
Bovis Construction Ltd
Costain Group plc
Y.J. Lovell (Holdings) Ltd
Norwest Holst Group Ltd
Tarmac Construction Ltd
Taylor Woodrow Construction Ltd
Wimpey Group Services Ltd

Thanks also go to the Construction Industry Training Board for the use of their safety training material; and, most importantly of all, to the staff of the Health and Safety Executive for their enthusiastic interest in our work and their considerable moral and financial support.

This work was funded by the UK's Health & Safety Executive. The contents of this article, including any opinions expressed, are those of the authors alone and do not necessarily reflect HSE policy.

References

Algera, J.A. (1990) Feedback system in organizations. In Cooper, C.L. and I.T. Robertson (eds), *International Review of Industrial and Organizational Psychology 1990*, **5**, pp. 169–93, John Wiley & Sons Ltd., London.

Arscott, P. (1980) *An Employer's Guide to Health and Safety Management*, Kogan Page, London.

Bass, B.M., W.F. Cascio and E.J. O'Connor (1987) Magnitude estimations of expressions of frequency and amount, *Journal of Applied Psychology*, **59** (3), pp. 313–20.

Bathgate, E.E. (1991) Field study to determine the preferred method of graphical feedback, BSc project, Department of Building Engineering, UMIST, Manchester.

Chokkar, J.S. and J.A. Wallin (1984) Improving safety through applied behaviour analysis, *Journal of Safety Research*, **15**, pp. 141–51.

Cohen, J. (1987) *Statistical Power Analysis for the Behaviour Sciences* (2nd edition). NY, Academic Press.

Cooper, M.D., R.A. Phillips, I.T. Robertson and A.R. Duff (1991) Improving safety on construction sites by the utilization of psychologically based techniques: Alternative approaches to the measurement of safety behaviour. *Paper presented at the 5th European Congress of the Psychology of Work and Organisation*, Rouen, France.

Dalton, A. (1989) Hard Labour – Hazards faced by building site workers. Television broadcast and programme support literature – *Dangerous Lives*, Channel Four, 06/09/89.

Davies, V.J. and K. Tomasin (1990) *Construction Safety Handbook*. London, Thomas Telford Ltd.

Guzzo, R.A., R.D. Jette and R.A. Katzell (1985) The effects of psychologically based interventions on worker productivity, *Personnel Psychology*, **38**, pp. 275–91.

Hale, A.R. (1984) Is safety training worthwhile? *Journal of Occupational Accidents*, **6**, pp. 17–33.

Health and Safety Executive (1988) *Blackspot Construction*. Health and Safety Executive (Library and Information Services), Sheffield.

Heinrich, H.W. (1959) *Industrial Accidents Prevention*. McGraw-Hill, New York.

Hollenbeck, J.R., C.R. Williams and H.J. Klein (1989) An empirical examination of the antecedents of commitment to difficult goals, *Journal of Applied Psychology*, **74** (1), pp. 18–23.

Hunter, J.E. and F.L. Schmidt (1990) *Meta-analysis*. Sage, London.

Kay, H. (1978) Accidents: Some facts and theories. *Psychology at Work*, (War, P.B. ed), Penguin, Harmondsworth.

Komaki, J. (1977) Alternative evaluation strategies in work settings: Reversal and multiple-baseline designs, *Journal of Organizational Behaviour Management*, **1**, pp. 53–77.

Komaki, J., K.D. Barwick and L.R. Scott (1978) A behavioural approach to occupational safety: Pinpointing and reinforcing safe performance in a food manufacturing plant, *Journal of Applied Psychology*, **63** (4), pp. 434–45.

Kopelman, R.E. (1980) Objective feedback. In Locke, E.A. (ed) *Generalizing from Laboratory to Field Setting*, Lexington Books, Lexington, MA.

Latham, G.P., M. Erez and E.A. Locke (1988) Resolving scientific disputes by the joint design of crucial experiments by the antagonists: Application to the Erez-Latham dispute regarding participation in goal-setting, *Journal of Applied Psychology*, **73**, pp. 753–72.

Copyright © 2001. All Rights Reseved.

Locke, E.A. and G.P. Latham (1990) *A Theory of Goal-setting and Task Performance*. Prentice-Hall Int (UK), London.

McAfee, R.B. and A.R. Winn (1989) The use of incentives/feedback to enhance work place safety: A critique of the literature, *Journal of Safety Research*, **20**, pp. 7–19.

O'Brien, R.M., A.M. Dickenson and M.P. Rosnow (1982) *Industrial Behaviour Modification*. Pergamon, New York.

Rhoton, W.W. (1980) A procedure to improve compliance with coal mine safety regulations, *Journal of Organizational Behaviour Management*, **4**, pp. 243–9.

Saarela, K.L., J. Saari and M. Aaltonen (1989) The effects of an information safety campaign in the shipbuilding industry, *Journal of Occupational Accidents*, **10**, pp. 255–66.

Shimmin, S., P.J. Leather and J. Wood (1981) *Attitudes and Behaviour about Safety on Construction Work*. Report to the Building Research Establishment by the Department of Behaviour in Organisations, University of Lancaster.

Whitener, E.M. (1990) Confusion of confidence intervals and credibility intervals in meta-analysis, *Journal of Applied Psychology*, **75**, pp. 315–21.

Wilson, H.G. (1989) Organizational behaviour and safety management in the construction industry, *Construction Management and Economics*, **7**, pp. 303–19.

Wood, R.E., A.J. Mento and E.A. Locke (1987) Task complexity as a moderator of goal effects: A meta-analysis, *Journal of Applied Psychology*, **72**, pp. 416–25.

Zohar, D. and N. Fussfeld (1981) A systems approach to organisational behaviour modification: Theoretical considerations and empirical evidence, *International Review of Applied Psychology*, **30**, pp. 491–505.

Copyright © 2001. All Rights Reseved.