SAFETY SCIENCE Monitor

ISSUE 2 1998

Article 3

VOL 2

EVALUATING THE EFFECTIVENESS OF INTERVENTIONS FOR OCCUPATIONAL TRAUMA

WORKSHOP ON OCCUPATIONAL TRAUMA - MEASUREMENT, INTERVENTION AND CONTROL, DECEMBER 14 - 16, 1998, BRUSSELS, BELGIUM

HARRY SHANNON

Professor, Clinical Epidemiology & Biostatistics, McMaster University, Hamilton, Ontario, Canada; and Senior Scientist, Institute for Work & Health, Toronto, Ontario, Canada

EVALUATING THE EFFECTIVENESS OF INTERVENTIONS FOR OCCUPATIONAL TRAUMA

There is often a divide between those whose job is to promote and put into place safety interventions (people such as policy makers or company managers) and those with the expertise and interest in evaluating how effective such interventions are (people such as researchers). I will discuss some possible reasons for this later. What I want to do in this talk is begin to bridge the gap, while recognizing the difficulties to be overcome. Indeed, I raise questions as much as I answer any. I will start by very briefly describing the context in which I work and in which I have most experience. I then consider what is known about factors related to better accident performance, and how to choose appropriate interventions. I will describe the need for evaluating interventions, and then consider the quality of evaluations. This will set the scene for further discussion of the issues involved using real life examples.

Occupational Health and Safety in Ontario

Paavo Kivisto in his paper has given an excellent description of the situation in Ontario. He has noted that the newly named Workplace Safety and Insurance Board is now responsible not only for workers' compensation, but also for injury (and illness) prevention. As well, the philosophy underlying the provincial legislation on Occupation and Health and Safety (OHS) is that of Internal Responsibility. That is, workplace parties should be able to resolve OHS issues internally, and only if this fails, should external involvement be needed. This approach fits with the general thrust to reduce red tape and the role of government. Several Safe Workplace Associations, specializing in various sectors of the economy, assist companies to improve their performance in OHS.

Despite the range of activities in Ontario, I think it is fair to say that a culture of evaluating the OHS activities has not existed. I think this is changing, and Paavo has pointed out the new emphasis on outcomes rather than process, which means the need to educate practitioners and policy-makers about what constitutes a good evaluation is acute.

Correlates of Better Accident Performance

The "Gold Standard" for determining if interventions are effective is the true experiment. That is, experimental units (individuals, work groups, etc.) are randomly allocated to receive either the new

approach or a sham (or standard) intervention. Unfortunately, evaluations of this nature are rarely if ever carried out in workplace safety, in part because of logistic difficulties. Rather, a variety of weaker study designs are conducted. The inherent nature of these non-experimental approaches means we can have less confidence in the validity of their conclusions. Yet even such evaluations are comparatively rare. For example, Johnston, Cattledge and Collins (1994) identified 3028 papers describing safety training. Of these, 198 reported some kind of evaluation, with just 41 (slightly over 1%) doing it well enough to meet their eligibility criteria. Many attempts to improve safety are treated as if they must work, because "common sense" says so, with little challenge to this view.

Instead, evidence used to demonstrate the value of the intervention typically comes from observational studies showing relationships between particular variables or characteristics that are correlated with higher or lower accident rates. Based on these relationships, interventions are developed to remove or modify conditions related to higher rates.

In many instances, the new measures can indeed be effective. However, in an era of limited resources, we can and should do better. Anyone who has taken an introductory course in statistics or research methods has been warned repeatedly that "correlations does not equal causation". Yet we commonly forget this truism when it comes to taking action. A simple, if rather extreme, example for another field can show the logical flaw in this. Suppose that 50 years ago, researchers studying lung cancer had not investigated the potential role of smoking, but rather examined physical characteristics of study subjects. They might well have noticed that those who developed lung cancer also had yellowing of their fingers, especially between their first and middle fingers. An obvious preventive measure to recommend would have been careful cleaning of the fingers. Of course, today we find this a ridiculous notion. It is "obvious" that the real problem is smoking. Yet Sir Richard Doll (1998), who conducted early important studies of smoking and lung cancer, recently wrote: "In retrospect it can now be seen that medical evidence of the harm done by smoking has been accumulating for 200 years". Indeed, referring to the initial skepticism in the 1950s about the role of cigarette smoking in lung cancer, he noted that those who were dubious "accepted that an association had been shown, but not that the association implied cause and effect." When we do not wish to accept a relationship, it is quite easy to explain it away.

In the safety field, though, the problem is more that we are too ready to accept relationships as causal. Quite apart from the "correlation versus causation" concern, there are often issues regarding the methodology of any type of study. I will later discuss these issues for studies evaluating safety interventions - for now let me just mention one - the quality of the data that are obtained on each accident, and how they are collected and classified. Even though it is typically recommended that "root causes" of accident be uncovered, investigations may be fairly superficial and attempt to place blame on particular individuals. Thus a supervisor conducting an investigation is unlikely to consider him/ herself at fault, but may attribute the accident to the carelessness of the injured worker. (In a study my colleagues and I conducted several years ago, our survey asked management respondents what they thought were the three main causes of occupational accidents. One of the replies was "number one - carelessness; number two - carelessness; number three - carelessness.") Clearly different approaches to understanding the causes of the accidents can lead to very different assessments of those causes, and ultimately to very different proposals for prevention.

Choosing appropriate interventions

Having identified (what we believe to be) causal factors, we must decide what intervention to apply in a given setting. A variety of factors will influence our choice, and in this section I will raise some of these.

An initial question asks: What is the burden of accidents/injury? If we do decide to target specific accidents, we need to think about where we will get the "biggest bang for the buck"; that is, where we expect the greatest cost-effectiveness. This is not necessarily as simple as seeing which accidents are the most common. Firstly, we may want to focus on accidents resulting in more serious injuries. The most common of these can be different from the most frequent accidents causing minor injuries. Secondly, as Davies, Stevens and Manning (1998) point out "simple classifications normally used for accident prevention ... create the impression that many accidents are similar" yet identical accidents are not common - which is why accident prevention is hard. Nevertheless, we can define criteria to choose the accidents we will try to prevent.

Having chosen the problem, in looking for solutions we do not want to re-invent the wheel. There may be proven solutions that we can use. Alas, the research literature contains relatively few. Just a few years ago, Stephen Guastello (1993) wrote a paper in Safety Science entitled "Do we really know how well our occupational accident prevention programs work?" His literature search turned up a disappointingly small number of evaluations - indeed remarkably few given the amount of time and effort devoted to occupational safety. As well, even when there are interventions that have been shown to be successful, knowledge and application of them is not always widespread. A report from the Institute of Medicine in the United States (1999) refers to "a yawning gap between what we already know about preventing or ameliorating injuries and what is being done in our communities, workplaces, and clinics".

The choice of a successful intervention also depends on the contextual or cultural factors at the workplace. What works in Scandinavia may not suit North American workplaces. For example, it appears that the former are far more open to "participatory approaches" whereas in North America there is much more likely to be insistence on "management prerogatives", and less involvement of the workforce in decision making. I know of one example in which the labour representatives on a committee were reluctant to get involved in decision making on safety - rather they behaved as passive "observers". Unions, just like management, have 'biases' regarding what types of intervention they dislike and what they prefer. The November 1998 issue of the Health, Safety and Environment Newsletter from the Canadian auto workers union criticises prevention based on accident proneness, behaviour based safety systems and experience rating in workers' compensation. There are also considerable differences between industries (even within jurisdictions), and I have seen differences in culture between two plants of a company within a single complex. Local knowledge will determine if an approach that is possible at one site may be totally unacceptable in another. This also applies, of course, to legislation in different jurisdictions. The prevailing political climate also plays a role, just as do fashions (some would say fads) in management.

Indeed, over time, the nature of safety measures have changed. Hale and Hovden (1998) have noted that the broad "scientific" approaches to safety have experienced three stages. The first stage, they point out, was concerned with technical measures - machine guarding, prevention of explosions and so on. This began in the nineteenth century and continued until after World War II. Accidents without technical causes were considered unpreventable. The second stage, which initially ran in parallel to the first stage, began between the world wars. Research was conducted into personnel selection, training and motivation as preventive measures. The two approaches finally merged in the 1960s and 1970s, via probabilistic risk analysis and ergonomics, becoming what is called human factors. Thus Menckel and Carter (1985) wrote: "Ideally, safety efforts should encompass both behavioural and technical aspects of the workplace …". The third stage focuses on management systems. It followed from "an increasing dissatisfaction with the idea that health and safety could be captured simply by matching the individual to technology." Hale and Hovden note that Heinrich's text of 1931 was one of the earliest expositions of this approach, "but the management principles which he expounded were largely unsupported by specific field research …". They note that "auditing approaches ... were also codified collections of practical expertise and judgement from people with long years of work in industry." I do not want to dismiss the wisdom and perception of such practitioners. But, disturbingly, Hale and Hovden reported that these approaches "had, and still have, little or no scientific basis".

Just as we should consider the frequency and severity of different accidents to maximize cost effectiveness, we also need to examine the level at which we intervene to greatest effect. Is it the worker, the work group, the workplace, the industry, or the jurisdiction? And even if we believe that the best approach is at the worker, work group, or workplace level, there may still be a role for policy makers to provide incentives through legislation or compensation systems. It certainly does seem to be very difficult to change individual behaviour in the long term - this is a finding common not only to occupational safety but many other spheres of activity. On the other hand, for example, there is reasonable success with safety incentive programs. Anecdotally, though, there are questions about whether they simply reduce reporting, rather than improve safety; and their effects may be relatively short term.

Other levels are those of the work group, the workplace and the jurisdiction. A little while ago I used the data from Stephen Guastello's paper mentioned above as well as a number of additional evaluations. In consultation with him, I divided the interventions according to the level at which they were aimed. Those aimed at selecting employees (based on such things as psychological tests) seemed to have little or no

effect. While the analysis was (necessarily) fairly crude, there was reasonable support for actions at the workplace level (Table 1).

Table 1
SUMMARY OF EFFECT SIZES FOR INTERVENTIONS OF DIFFERENT TYPES
(after Guastello, 1993)

INTERVENTION LEVEL AND TYPE	NUMBER OF STUDIES	AVERAGE EFFECT SIZE[1] (%)
Aimed at the individual:		
Personnel selection	30	3
Exercise and stress management	2	15
Behavior modification	8	54[2]
Targeted skills training	1	30
Information-based campaigns	1	-33[3]
Aimed at the group:		
Quality circles/discussion groups	2	33
Autonomous work groups	1	(160)[4]
Aimed at the workplace:		
Management procedures	3	49
Housekeeping	1	(50-75)
International Safety Rating System	4	17
Safety committees	6	24
Comprehensive ergonomics	4	46
Cooperative compliance program	7	46
Aimed at the jurisdiction:		
OSHA (US) - Global	7	3
- Specific	12	19
Finnish national program	2	18
Brazilian legislation	1	70[5]

[1]Based on comparison of reduction of accident rate in intervention group relative to reduction in control group; or on correlation when applicable. See Guastello (1993) for details. Note that quality of studies not incorporated into analysis.
[2]With exceptional case removed, average effect size was 35%.
[3]Poorer performance in intervention group than in control group.
[4]Value exceeds 100 because of increase in accidents in control group.
[5]Reduction claimed by authors to be due to legislation, but observations made over considerable period of time.

The final item in determining the nature of an intervention is whether the goal is short-term or long-term. In some ways, the question seems ridiculous. Who could possibly argue that we should only attempt to fulfil short-term aims? Yet there are legitimate reasons for short-term aims. Especially with a participatory approach, not only may the goals of the program have a long time horizon, but the time to

implement such an intervention may also be long - as I am sure many of you have experienced. The workplace parties will quite rightly want to make some improvements more quickly, yet long-term maintenance of, e.g., behaviour change can be a challenge.

Need for evaluation of interventions

Researchers often call for evaluation of safety interventions. Yet good published evaluations are rare. There are a number of reasons for this - for example, the culture is typically not one that encourages such evaluations. Governments, for example, may have ideological or political reasons for taking action; or those who have struggled long and hard to persuade their organization to undertake some program do not want to open up the possibility that the program is ineffective - thereby reducing their own credibility and ability to implement any future proposals. Yet, particularly with educational or other behavioural change programs, there can be unintended harmful effects detectable only through well designed research. As Robertson (1992: 116) points out, "[a] major barrier to the scientific evaluation of programs is the reluctance of those who develop, advocate, or profit from programs to have them evaluated objectively." As well, it is difficult to remove established programs even if they are found to be ineffective or harmful.

So why do evaluations? Firstly, some "obviously" sensible interventions simply don't work - or can even be dangerous. Secondly, when interventions are found to be unsuccessful, publication of this fact should prevent the wasted time and resources that others might use if they try the same program. (Of course, the program may be more applicable in their situation, or applied in a somewhat different fashion, but that is a separate issue.) Notice that I added that publication of the results is needed. Sometimes, companies may consider such information to be proprietary and part of their competitive advantage. One way around this would be for policy makers to provide incentives for companies willing to do (or get others to do) careful evaluations and make the results widely available.

An example can help show why, even if an observed relationship is indeed causal, we need to determine if the intervention we propose is appropriate: An almost universally acknowledged factor in safety is "management commitment". One illustration of this may be attendance by the senior manager at health and safety committee meetings. An obvious recommendation would be that the senior manager of every company attend such meetings. This might not prove effective for the following reason: the relationship observed between commitment and safety may have been present because those managers genuinely concerned about safety showed this by their actions and attitude. Yet if managers without such concern are pressured into attending committee meetings, their lack of interest will likely become clear to other committee members. They will be anxious to end meetings quickly, or may be called out of meeting to deal with "more important" matters. Their attitude could even do more harm than good by interfering with the functioning of the committee. We can only check if this happens through a proper evaluation.

I do not want to downplay the difficulties that can prevent or accompany evaluations. Wickström and colleagues note that these days there is continuous change in the organization of work and in work tasks. If safety interventions are superimposed on such situations, it is hard, perhaps impossible, to separate the relative effects of the changes. Wickström et al. also noted that they had begun their intervention at a second site, a shipyard. The company went bankrupt, so that component of the study had to be abandoned.

I will conclude this section with a question for discussion. It arises from my earlier comment that many organizations do not have a culture that encourages evaluation. Such organizations may be put off by the prospect of a detailed comprehensive evaluation which would use some of their time and resources, even if done by outside researchers. In such circumstances, is a poor evaluation better than no evaluation? The main potential benefit is that it can begin to create a climate in which good careful evaluations become the norm. Yet a serious concern is that a poor evaluation could reach the wrong conclusion about the particular program being assessed.

Judging the quality of evaluations

A lot has been written about how to do good evaluations. However, none of the books and few papers that I have seen apply the methodologies specifically to occupational safety interventions. (Recent papers by Zwerling et al. (1997), Goldenhar and Schulte (1994, 1996), and Schulte, Goldenhar and Connally (1996) are notable discussions of this topic.) Because of this, my colleagues and I have written a paper to be published soon in Safety Science that describes "how to read" such an evaluation. We are

currently developing a manual on "how to do" one. I should note that we have a good deal of assistance in this effort. Firstly, members of SCOAP have been generous with their feedback. As well, in the U.S. there is an effort to develop a National Occupational Research Agenda (NORA). One of their subgroups is examining the role of evaluation, and members of that group are also helping us.

We have identified several criteria determining the quality of an evaluation, and have grouped them into a number of categories.

1 - Program objectives and conceptual bases

It may seem obvious that the objectives of a program need to be stated in advance, but I have seen interventions in which written material outlining the aims of the program were unavailable. While the person(s) who developed the program may know, others may be at a loss to understand. As well, appreciating the theoretical rationale for the program is important - it helps interpretation of results, and is critical if the program is to be adapted to different situations.

2 - Study Design

I have already noted that the gold standard for evaluations (and many other questions) is the true experiment. In practice, this is difficult or impossible to achieve in the working environment. One option is to use "quasi-experimental designs" which mimic the randomized design as far as possible. Features of such designs include the presence of a control (comparison) group, and the use of outcome measures both before and after the interventions. Some other designs seen in the literature are unfortunately so weak that their results are best ignored. (Their real value may be, as I noted earlier, in promoting the climate in which evaluation of new programs is the norm.)

3 - External validity

The aim of a good study design is to ensure lack of bias within the study. This is known as internal validity. Those who read the paper may also wonder whether they could expect similar results if they applied the intervention in their own setting. As I noted earlier, contextual factors may influence how well a program works. The extent to which we can generalize from one setting to another is known as external validity. Researchers can help others to gauge external validity by providing a variety of information. This includes a description of the program participants as well as the intervention itself. It is also important to add contextual factors, such as local legislation and the presence or absence of a workers' compensation system.

4 - Outcome measurement

The outcome data we use may depend not only on the type of intervention, but also on the level at which it is aimed. When jurisdictional level changes are made (legislation or workers' compensation policy), the numbers of accidents are likely sufficient to detect changes in accident rates. As Paavo has noted, we will typically want to check rates of fatal accidents, lost time accidents and all injury accidents, as well as severity measures. When our interventions are targeted at the workplace level, however, the actual number of accidents is often relatively low. From a statistical viewpoint, we have little "power" to demonstrate an effect of the intervention. In such circumstances it is only feasible to detect a change in intermediate outcomes, such as safety knowledge or safety behaviour on the job (Figure 1). Ideally, we would like evidence from some other source to confirm that changes in these intermediate outcomes are indeed related to changes in accident rate. We have good reason to be cautious. As noted earlier, behavioural interventions can have unintended harmful consequences. (The option to use intermediate outcomes is not available when the intervention is a technical measure.)

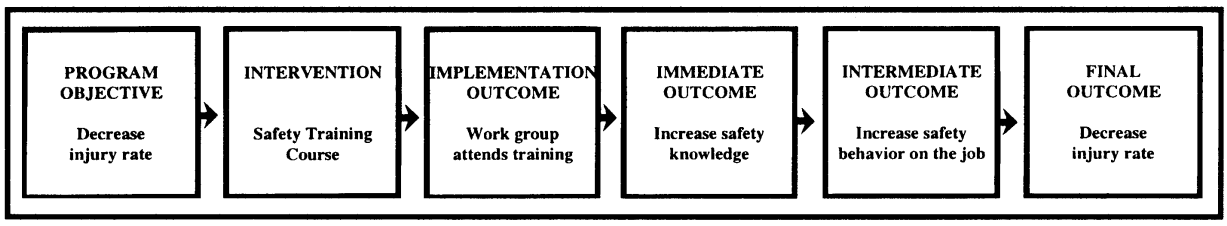

Fig. 1. Relationship between program objectives, intervention and outcomes.

When we can use accidents as the outcome, it is important that a rate is computed, rather than simply the absolute number of accidents. I have seen papers where the latter was done, without any indication as to whether the number of workers involved also changed. This, of course, is a critical piece of information.

Finally, our measurement methods must be shown to be valid and reliable. For example, accident reporting can influence accident rates. Thus, workers involved in a group safety incentive programme will experience peer pressure not to report accidents. With experience rating in workers' compensation, companies may also be reluctant to document claims. In Ontario, for example, consultants in claims management advise management on how to reduce the number and cost of claims. Note the phrasing - number of claims, not number of accidents. One trick I have heard of is to try to argue that claims for back pain should not be registered as "new" claims, but considered as "reopened" claims from a previous injury which occurred in a previous job. Thus the episode of back pain may be attributed to a previous employer, from the time of the initial claim. This alone could improve the apparent accident rate for the company. As well, measures of safety behaviour depend on who observes the behaviour. Observer variability can be reduced by providing strict criteria for assessments, and preparing observers carefully. Discussion of differences in their ratings by the observers during pilot testing is usually useful. Even so, some variability will remain - and examination of it can constitute whole sub-studies on their own.

5 - Qualitative data

In the last few years, there has been a gradual realization among those whose expertise is in the quantitative fields that a great deal can be learned from qualitative methods. I am not suggesting that qualitative approaches replace quantitative ones. What I would argue, though, is that careful analysis of transcripts of interviews with key informants can help explain why a program was (or was not) effective. Otherwise, many programs are effectively "black boxes" and we are left with little understanding of why the program worked out the way it did. This is particularly important when a program is apparently ineffective, since we may wish to modify some of its components in an effort to improve it. Comparable results from qualitative and quantitative data serve to strengthen the conclusions.

6 - Threats to internal validity

For reasons of space and time, I will not go into detail about aspects of a study that can comprise internal validity. I will simply point out that these threats to validity are particularly important in quasi or non-experimental designs - but they cannot be ignored even in randomized studies. The list includes such items as: differential attrition; contamination; non-compliance; selection effects; etc. Their importance is that they can make an intervention appear to be successful when it really isn't, or vice versa.

7 - Statistical analysis

I am often dismayed by the quality of statistical analysis in safety evaluations. I am not singling out safety practitioners and researchers. Statisticians have examined statistical methods used in a variety of fields, and generally found them wanting. Disturbingly, such reviews sometimes state that the conclusions of several papers are likely wrong and would be different if a proper analysis had been performed.

The errors can often be subtle and not obvious to most readers. I recently read a paper in which the authors used archival data to evaluate a particular program. They appropriately used a time series analysis. As well, they made an adjustment which, they claimed, allowed for the difference in average company size between the intervention and control groups. They concluded there was a strong beneficial effect of the program. Unfortunately, the adjustment for company size was not sufficient to account for the differences

and there is enough information in the paper to suggest that the effect of the program was, if anything, in favour of the control group! (The nature of the data available meant that the comparison was inherently biased in favour of the control group; had proper allowance for this bias been possible, it might have shown the program was effective, but almost certainly much less so than was claimed.)

Regarding specific statistical issues, I will mention only one specifically: sample size. It should really be considered at the design stage, before the study even begins. (Gregersen, Brehmer and Moren, 1996, provide a rare example.) When studies are small, it may be next to impossible to detect even a real and substantial effect of a program. This may even stop an evaluation from proceeding. As one epidemiologist has stated: one of the first decisions is whether the sample size should be zero or non-zero! (He was also arguing that a poor quality study may be worse than no study.) A reader of an evaluation can assess if the sample was large enough if a confidence interval is provided. When the study is too small, the confidence interval on the effect size will be very wide, showing that the estimate is very imprecise.

8 – Conclusions

We all want the programs we develop and/or implement to work, but they do not always do so. Nevertheless, it is tempting (even subconsciously) to put the best light on our failures; thus we may downplay just how ineffective our intervention was, or even over interpret our data. An example of this is a published study of training in five plants. While the authors noted the impossibility of singling out any one factor as the key variable determining program outcomes, they still listed four program implementation variables as being associated with better training outcomes. This is taking the data too far - the authors could not reasonably make this claim.

In drawing conclusions, we need to allow for the limitations in our study, and as I have emphasized most studies in safety evaluation have important methodological limitations. Authors need to play the "Devil's Advocate" in discussing their results, and seek ways to improve the quality of their evaluations. After all, none of us want to waste time, effort and resources on ineffective safety measures, yet we can get the "wrong answer" if our evaluations are poor.

As I noted earlier, the optimal study design is the randomized controlled trial (RCT). In medical research, there are several examples of specialties in which RCTs were claimed to be impossible, for ethical or other reasons. Yet eventually the RCT was introduced and became accepted as the standard of proof. Likewise, there are some examples of RCTs of safety interventions, but they are the exception, not the rule. While I am not naive about the barriers, I urge evaluators not to immediately dismiss the RCT in occupational safety research. Certainly, it will often be impossible to conduct one - but if we keep in mind this approach, opportunities should appear.

Examples of Successful and Unsuccessful Interventions

I hope this paper has provided the basis for a broad discussion on real life examples of evaluations of safety interventions and why the interventions were successful or unsuccessful. Good quality studies are perhaps the best means we have to truly improve health and safety at work.

Acknowledgement

I thank Lynda Robson for her useful suggestions and comments on a draft of this paper.

REFERENCES

Bonnie, R.J., Fulco, C.E., Liverman, C.T., Eds (1999) *Reducing the burden of injury advancing prevention and treatment.* Committee on Injury Prevention and Control Division of Health Promotion and Disease Prevention, Institute of Medicine, National Academy Press, Washington, D.C. "Quotation from Summary available on Internet (November 1998) at http://www.nap.edu/readingroom/books/burden/".

Doll, R. Uncovering the effects of smoking: historical perspective. (1998) *Statistical Methods in Medical Research* ***7****, 87-117.*

Davies, J.C., Stevens, G., Manning, D.P. (1998) Understanding accident mechanisms: an analysis of the components of 2516 accidents collected in a MAIM database. *Safety Science* **29**, 25-58.

Goldenhar, L.M. and Schulte, P.A. (1994) Intervention research in occupational health and safety. *Journal of Occupational Medicine* **36**, 763-775.

Goldenhar, L.M. and Schulte, P.A. (1996) Methodological issues for intervention research in occupational health and safety. *American Journal of Industrial Medicine* **29**, 289-294.

Gregersen N.P., Brehmer, B. and Moren, B. (1996) Road safety improvement in large companies. An experimental comparison of different measures. *Accident Analysis and Prevention* **28**, 297-306.

Guastello S.J. (1993) Do we really know how well our occupational accident prevention programs work? *Safety Science* **16**, 445-463.

Hale A.R. and Hovden J. (1998) Management and culture: the third age of safety. A review of approaches to organisational aspects of safety health and environment. In *Occupational Injury: Risk, Prevention & Injury,* eds A. Williamson and A.-M. Feyer, Taylor & Francis.

Johnston, J.J., Cattledge, G.T.H. and Collins, J.W. (1994) The efficacy of training for occupational injury control. *Occupational Medicine: State of the Art Reviews* **9**, 147-158.

Menckel, E., Carter, N. (1985) The development and evaluation of accident prevention routines: a case study. *Journal of Safety Research: **16**, 73-82.*

Moran, G.E. (1985) Regulatory strategies for workplace injury reduction. *Evaluation Review* **9**, 21-33.

Robertson, L.S. *Injury Epidemiology.* (1992) Oxford University Press

Schulte, P.A., Goldenhar, L.M. and Connally, L.B. (1996) Intervention research: science, skills, and strategies. *American Journal of Industrial Medicine* **29**, 285-288.

Zwerling, C., Daltroy, L.H., Fine, L.J., Johnston, J.J., Melius, J. and Silverstein, B.A. (1997) Design and conduct of occupational injury intervention studies: a review of evaluation strategies. *American Journal of Industrial Medicine* **32**, 164-179.

APPENDIX A

THEMES FOR BRUSSELS WORKSHOP

1. Correlates of better accident performance

A number of workplace characteristics have been shown to be related to better accident records. However, there are several issues to be considered in transferring such information to intervention programmes.

a) Are the relationships truly causal?

b) What are good performance measures? Are they valid and reliable?

c) How good are the data obtained on each accident, and how are the data collected and classified?

2. Choosing appropriate interventions

Having identified causal factors, we must decide what intervention to apply in a given setting. How do we do so?

a) What is known to work? (Prior studies, etc)

b) What is the ”burden of injury”, i.e. what are the most common types of accident / injury?

c) How do contextual / cultural factors affect the choice? What works in Scandinavia may not suit North American workplaces; or different industries may not be open to certain approaches.

d) Accidents are each different, making prevention hard. How can we deal with this?

e) At what level do we intervene to greatest effect? The worker, the work group, the workplace, the industry, or the jurisdiction?

f) Is our goal short-term or long-term?

3. Need for evaluation of interventions

Researchers often call for evaluation of safety interventions. But action-oriented organisations often do not have a culture that encourages such evaluations. So why do them?

a) Many ”obviously” sensible interventions don’t work.

b) Interventions found to be unsuccessful won’t be tried elsewhere.

4. How can you tell if an evaluation is well done? And how can you do a good evaluation?

List of criteria - they are in a paper to be published soon in Safety Science. They describe ”how to read” an evaluation; we are also writing a manual on ”how to do” one.

5. Examples of Successful and Unsuccessful Interventions

Published examples of interventions that ”worked” and some that didn’t.