On the difficulty of promoting workers' safety behaviour: overcoming the underweighting of routine risks

Dov Zohar*

Faculty of Management, Technion Institute of Technology,
Haifa 32000, Israel and the Institute for Work and Health,
481 University Avenue, Toronto, M5G 2E9, Canada
E-mail: dzohar@tx.technion.ac.il
*Corresponding author

Ido Erev

Faculty of Management, Technion Institute of Technology,
Haifa 32000, Israel
E-mail: erev@tx.technion.ac.il

Abstract: The paper seeks to explain the seeming paradox whereby workers' safety behaviour during routine work depends largely on supervisory contingencies (i.e. pressures and rewards), rather than on self-preservation. We identify three behavioural tendencies accountable for underweighting of outcomes associated with safe behaviour, i.e. delayed outcomes (i.e. melioration bias), rare or uncertain outcomes (i.e. recency bias), and outcomes concerning social externalities. Jointly, they result in a tendency to favour unsafe behaviour in many routine work situations. Examination of effective and ineffective intervention programmes suggests that the key to success lies in providing frequent, personally meaningful, and immediate rewards for safe conduct, overriding the costs associated with that behaviour and exceeding the benefits of unsafe behaviour.

Keywords: behavioural safety; risk perception; safety climate.

Reference to this paper should be made as follows: Zohar, D. and Erev, I. (2007) 'On the difficulty of promoting workers' safety behaviour: overcoming the underweighting of routine risks', *Int. J. Risk Assessment and Management*, Vol. 7, No. 2, pp.122–136.

Biographical notes: Dov Zohar got his PhD in Psychology from the University of Maryland in 1976. He is an Associate Professor at the Faculty of Industrial Engineering and Management, and a senior scientist at the Center for Work Safety at the Technion. His research includes the study of organisational climate and culture, in particular safety climate in high-risk organisations, as well as the study of climate-leadership relationships. His research covers climate-measurement and climate-improvement issues, focusing on the key role of strategic and supervisory leadership.

Ido Erev got his PhD in Psychology from the University of North Carolina in 1990. He is a Professor of Industrial Engineering and Management, and a member of the Max Wertheimer Minerva Center for Cognitive Research, and of the Center for Work Safety at the Technion. His research focuses on

Copyright © 2007 Inderscience Enterprises Ltd.

modeling and deriving the implications of human adaptation to economic incentives. In his current applied research he focuses on the design of gentle rule enforcement systems and training methods.

1 Introduction

Safety behaviour presents a paradox to practitioners and researchers alike because, contrary to the assumption that self-preservation overrides other motives (Maslow, 1970), careless behaviour prevails during many routine jobs, making safe behaviour an ongoing managerial challenge. This is evident in the fact that failure to use the protective gear *provided at the workplace* accounts for about 40% of work accidents and occupational disease. This statistic has not changed in more than 20 years despite ongoing efforts (National Safety Council, 1999). More direct evidence comes from a recent study in which systematic behaviour observations by trained observers in 423 workgroups across industrial sectors indicated an incidence rate of 33.8% for unsafe behaviour, using criteria provided by line managers and safety officials in each company (Zohar and Luria, 2005). Unsafe behaviour associated with potentially disastrous consequences (e.g. plant-wide damage due to chemical or fire hazards) was less prevalent, though quite high, averaging 21.7%. Clearly, therefore, safe behaviour in routine work poses managerial challenges. Why is it so difficult to promote safe behaviour on shop floors? Why is workers' behaviour careless despite obvious risks on the work floor? Drawing on known behavioural biases, we first offer a decision-making analysis of this phenomenon. Subsequently, we examine some major intervention strategies and suggest that their effectiveness depends on the extent to which these biases are taken into account.

2 A decision-making perspective

The decision-making literature highlights three behavioural biases or tendencies that would lead individuals to violate safety rules designed to protect them. All three tendencies suggest that people are likely to behave as if they underweight possible outcomes of unsafe behaviour, so that the appraised benefits of unsafe behaviour will often outweigh those of safe behaviour. The frequent and immediate benefits of unsafe behaviour, such as faster pace, lesser effort, or greater personal comfort, are thus given greater weight than low-probability dangers that could be averted by safer behaviour. Furthermore, given the low probabilities involved, both safe and unsafe behaviour result in *non-events* most of the time (i.e. avoidance of low-probability injury), highlighting the benefits of unsafe behaviour. As a result, the subjective expected utility of unsafe behaviour exceeds that of safe behaviour in many routine work situations. Consequently, choosing safe behaviour in routine work is not as natural as one would assume based on the self-preservation postulate.

2.1 Underweighting delayed outcomes

The first behavioural bias we discuss relates to the immediacy of decision outcomes. Violations of safety rules associated with routine work often lead to immediate and assured positive outcomes (higher speed, greater comfort) as well as delayed and unsure (i.e. probabilistic) negative outcomes associated with accidents or near-misses. In some cases, negative outcomes such as hearing loss or carcinogenic disease following chronic exposure are substantially delayed yet virtually assured. Thus, violating a rule that requires the use of ear protectors in a noisy workplace saves discomfort and may also improve performance through better detection of auditory signals, but it also leads to cumulative hearing loss which becomes evident some ten or 15 years later. The extensive study of animal and human behaviour conducted by Herrnstein, Loewenstein and their associates, see (review in Herrnstein et al., 1993), reveals that decision makers tend to underweight delayed outcomes even if they are highly likely or virtually assured. This phenomenon, known as melioration, implies violations of safety rules (e.g. rules concerning ear- or respiratory protection) even under conditions where such rules maximise utility in the long run due to high negative utility associated with the delayed outcomes.

In a typical melioration study (e.g. Herrnstein et al., 1993), participants play a simple computer game. In each round of this game the subjects are asked to choose between two keys – Left or Right. Pressing the Left key causes a coin to fall from a box on the left side of the screen, and the Right key causes a coin to fall from a box on the right side of the screen. Participants are told that the coins they collect will belong to them and will determine their payoff. In one of the games run by Herrnstein et al. (Herrnstein et al., 1993) all coins had the same value. The subjects played the game for 15 minutes, and had to minimise coins' falling time in order to maximise earnings. The coins that dropped from the right box took 2 seconds less to fall than coins from the left box. However, each Right choice increased the falling time of each of the next ten coins by 0.4 second. Thus, each Right choice would save 2 seconds in the short term, but cost 4 additional seconds over the long term. Thus, the net loss of a Right choice was 2 seconds. Since subjects had limited time to collect coins (15 minutes), loss of time was equivalent to a loss of money. The optimal utility-maximising behaviour would be to choose the Left key on every trial (except for the last six), thus minimising the time needed for a coin to drop. Melioration dictates choosing the Right key each time, which is maximal inefficiency. This behaviour follows the intuition of pressing the key for the coins that drop most rapidly *at the moment*, ignoring how the choice affects the falling times of later coin drops. Overall, 79% of the choices in the second half of this experimental condition reflected melioration.

Other studies reveal that the above tendency to underweight delayed outcomes is quite robust, i.e. people frequently underweight (discount) future earnings even when the distribution of outcomes is well-known, resulting in discounts exceeding rational economic considerations; see the review by Frederick et al. (2002). Herrnstein et al. (1993) speculate that many examples of suboptimal behaviour are activated by melioration. One interesting example is the workaholic syndrome that prevails among professionals who can control their pace and amount of work (Harpaz and Snir, 2003). Presumably, workaholics overweigh immediate task outcomes and are less sensitive to the possibility of long-term negative implications that might affect social relationships and family roles. Under-saving for retirement in order to maintain current higher standard of living is another example (Thaler and Shefrin, 1981). In this case, immediate and frequent rewards are given greater weight than hardships which may be encountered some 30 years later. Both these examples can be compared to situations where workers opt for unsafe behaviour by underweighting long-term consequences of failure to use personal protective equipment in favour of short-term gains. Thus, melioration bias provides an explanation for the prevalence of unsafe behaviour in routine work.

2.2 Underweighting low probability (rare) outcomes

Another behavioural bias relates to the fact that negative outcomes of violating safety rules occur only rarely, with safe (or unsafe) behaviour resulting most of the time in non-events (i.e. no injury). In a landmark study, Heinrich (1931) found that only one out of 15,000 safety violations leads to an accident. More recent estimates put this figure even higher, suggesting that 30,000 violations (unsafe acts and/or unsafe conditions) will result in only 30 minor injuries or one fatality (Reason, 1997). This ratio encourages safety-rule violations because people tend to overweight recent outcomes when choosing among action alternatives, otherwise known as recency bias (Barron and Erev, 2003). Recurring, hence recent benefits of unsafe behaviour are thus overweighed by comparison to rare (hence non-recent) events associated with injury or near-miss, resulting in underweighting of these events (Barkan et al., 1998; Barron and Erev, 2003; Perry et al., 2002).

Laboratory studies that demonstrate this behavioural tendency have used subjects whose task is to press one of two unmarked buttons on a computer screen over a series of many trials. One button is associated with a fixed payoff while the other is associated with variable payoff whose value is randomly selected from an underlying distribution after each trial. The outcome is presented to the decision maker immediately after each choice, helping him/her learn the underlying payoff distributions for the buttons as the experimental session continues. In one of the conditions, Barron and Erev (2003) gave the participants 4000 Israeli pennies (about $10) as a show-up fee, instructing them that they would lose some of them in the experiment, which lasted for 400 trials. In this study, the selection of one button (S) yielded a loss of three pennies, and selection of the other button (R) yielded a loss of 32 pennies in 10% of the trials, and 0 in 90% of the trials. Subjects were thus presented with the following choice situation:

Problem 1 Choose one of two alternatives (the percentage on the right presents the rate of S choices):

S (safe) get −3 with certainty [40%]
R (risky) get −32 with p = 0.1; 0 otherwise (0 = 0.9).

Although the safe option (S) maximised earning, the modal response was R. The rate of S choice over the 400 trials was 40%, remaining relatively stable across trials.

Other choice problems incorporated in this study suggest that the tendency to prefer the risky option in Problem 1 should not be interpreted as indication of risk seeking in the loss domain. Thus, in Problem 2 most participants preferred S, although R offered higher payoff in the long run:

Problem 2 Choose between:

S (safe) get −9 with certainty [63%]
R (risky) get −10 with p = 0.9; 0 otherwise (p = 0.1).

See similar results in Perry et al. (2002) and Yechiam et al. (2002). Barron and Erev (2003) explain this pattern with the proposition that decision makers *underweight* small-probability outcomes (e.g. the −32 outcome in Problem 1, and the 0 outcome in Problem 2). Analysis of this proposition suggests that a tendency to underweight rare events can emerge from an adaptation or learning process that overweights recent outcomes. To see why a recency effect implies underweighting of rare events consider a decision maker who presents strong recency effect in Problem 1. Each time, this individual selects the option that he/she found most effective at the previous decision point. Since R yields higher payoff in 90% of trials, this implies 90% R choices.

The tendency to underweight rare events in situations involving repeated decisions seems to be rather stable. Barkan et al. (1998) observed a similar pattern in categorisation decisions; Perry et al. (2002) observed this pattern when participants received precise description of decision problems, and Hertwig et al. (2004) observed a similar pattern when participants could sample and learn the payoff distribution[1]. Thus, recency bias provides another explanation for choosing unsafe behaviour over the safe alternative in routine safety dilemmas.

In an apparent contradiction to the results summarised above, studies of risk assessment provide evidence for overweighting of rare events. For example, Lichtenstein et al. (1978) observed a consistent overestimation of the probabilities related to the rarest causes of death. A similar finding is that teens greatly overweight the chances of death in the near future, estimating the probability to be 18.6% whereas the actual probability is 0.04% (Fischhoff et al., 2000). Another example is that when Americans were asked to estimate the probability that a smoker will develop lung cancer, the mean estimate was 38% while the actual probability is between 6 and 13% (Viscusi, 1992). Finally, risk perception studies at work have repeatedly shown that workers over-estimate the likelihood of injury (or near-miss) at the workplace while, at the same time, behaving as if this likelihood is much smaller (Greening, 1997; Reason, 1997). Thus, people over-estimate low probabilities, but behave (in repeated settings) as if they underweight the (rare) events associated with such small probabilities. In a recent paper Barron et al. (2004) show that the difference between estimation and behaviour is not driven by the difference between tasks. Such a difference can appear with regard to the same laboratory task when a person is asked to estimate probabilities and to make decisions.

2.3 *Underweighting social externalities*

When analysing decision making, it is important to distinguish between two types of outcome, i.e. 'internalities' and 'externalities'. Internalities are the outcomes that affect the decision maker. Externalities are outcomes that affect other individuals as a result of his/her decision. As noted above, in routine work situations the decision to ignore safety rules often has positive average internalities (e.g. saving time and effort by using shorter procedures). Social externalities, on the other hand, are often negative because safety shortcuts may create unsafe conditions that endanger other individuals working nearby, or increase the likelihood of an accident in which others could also be injured.

Situations in which the externalities are negative and have larger absolute values than positive internalities are known as 'social dilemmas'; see Dawes (1980) and Messick and Brewer (1983)[2]. In the context of safety, social dilemmas are social interactions characterised by two properties:

- each individual benefits (on average) from violating safety rules (e.g. gains speed, reduces discomfort), but
- each individual is better off if all group members observe the rules, thus creating a safer environment.

Thus, in routine work-settings a tendency to underweight social externalities implies 'rational' violations of socially desirable safety rules. To clarify the logic behind such rational violations we can consider a simple numerical example. Assume that each member of a three-member workgroup can select one of two methods (Safe and Risky) to perform a particular task:

S: Safe method (observing safety rules).

R: Risky method (saves five minutes but violates safety rules, which may lead to an accident affecting other team members).

If R is selected, there is a probability of 0.01 for injury to any group member. Cost of injury averages 300-minute loss. Thus, the expected cost of each R choice is 300 * (0.01) = three minutes per group member. Table 1 shows the expected outcome for Worker i as a function of his/her choice and the choice made by the other workers. The table indicates that R is a dominant strategy, offering the highest utility for each individual worker. Independent of the choice of the other two group members, Worker i always saves two minutes by selecting this strategy. Yet, if all group members follow their dominant strategy (loss of four minutes), each worker will lose four minutes as compared to when all workers behave safely (no loss). In other words, the internalities are +2, and externalities are −6.

Table 1 Player i's payoff in a numerical example of a safety dilemma as a function of choice and the number of other R choices in the group. The group includes three workers (i and two other members). Note that R (violation of the safety rule) is a dominant strategy (payoff from R is higher than payoff from S). Yet, when all workers select R, the payoff (-4) is lower than when all workers select S (0)

Worker i's choice	*# of other R choices*		
	0	1	2
R	+2	−1	−4
S	0	−3	−6

Typical experimental studies of choice behaviour in social dilemmas present the participants with a simplified 'investment' decision. For example, in one of the conditions summarised by Bornstein and Ben-Yossef (1994), subjects were divided into three-member teams. Each participant received an endowment equal to five Israeli Shekels (about $2 at the time of the experiment) and was asked to choose between 'investing' and 'not investing' it. Team members had to decide individually without communicating with each other, but their bonus was a function of the decisions made by all three. The bonus (in Shekels) was three times the number of investors in the group. In addition to this bonus, those who did *not* invest kept their five-Shekel endowment. Notice that this 'investment problem' is an economic version of the safety dilemma presented above. As in the safety dilemma, the not-investing alternative is dominant, but it impairs the group's payoff.

In most experiments, investment/cooperation rate is below 50%. In the study described above, observed investment rate was 24%. Thus, although the investment rate was higher than that expected from assuming that subjects 'free ride' and follow their dominant strategy (i.e. underweight externalities, resulting in 0% investment), it is much lower than the 100% investment that maximises group earnings. Other studies of social dilemmas show that the tendency to free-ride is rather robust and tends to increase with repetition[3]. An example of the significance of this dilemma in work environments is provided in a field study of orange picking (Erev et al., 1993). This study shows that a collective reward rule allowing free riding (i.e. equal division of a group-level bonus regardless of individual contribution) can decrease the productivity of individual members by as much as 50% or more.

In regard to safety dilemmas, free riding can also be a product of implicit decisions. Implicit free-riding possibilities occur, for example, when more than one worker can detect and repair hazardous technical problems. When the responsibility of detection is shared, but its cost is paid by the detector (e.g. extra time or effort), workers may be motivated to free-ride and let a co-worker detect and deal with the problem. Erev et al. (1995) demonstrated a free-riding tendency of this type. All of the above suggests that the tendency to choose *un*safe behaviour in many routine work situations remains unaltered even when such decisions entail negative externalities. This behavioural bias thus provides additional explanation for the prevalence of unsafe behaviour in routine work.

3 Effective interventions modify the value function of safety behaviour

Given that unsafe behaviour in many routine work situations is sustained by behavioural biases whose effects are probably cumulative, promotion of safety behaviour must overcome these biases by modifying the perceived value function (i.e. self-assigned weight) for safe behaviour. Our analysis suggests that frequent, short-term, personally meaningful rewards that outweigh perceived costs must be introduced.

The primary safety-intervention strategy that meets these criteria is known as 'behavioural safety' or 'behaviour-based safety' (Geller, 1996; McAfee and Winn, 1989; O'Hara et al., 1985). This strategy employs the organisational behaviour modification (OBM) protocol (Luthans and Kreitner, 1985), comprising identification of desired behaviours, observation-based measurement of behaviour frequency, and provision of contingent and immediate reinforcement aimed at performance improvement. Operationally, this results in provision of frequent, short-term, personally-meaningful rewards. This framework derives from Skinner's (Skinner, 1974) reinforcement protocols as well as other abstractions of human adjustment to the incentive structure, such as Thorndike's Law of Effect (Bandura, 1969, 1986; Erev and Roth, 1998; Thorndike, 1898). Behavioural safety interventions are likely to be effective whether workers are described as 'rational expected-utility maximisers' or 'adaptive learners' (i.e. rational agents discover optimal behaviour based on premeditated utility-maximising considerations whereas less rational agents are expected to exhibit similar behaviour as a result of an adaptive learning process).

Most behaviour-based safety interventions employ publicly displayed charts of the frequency (%) of focal behaviours as compared with designated targets. Data are collected by external observers or co-workers, and the frequently updated discrepancy between current level and designated goal provides the immediate incentive or reinforcement for change. In a recent example, Lingard and Rowlinson (1997) launched their intervention on construction sites with a series of joint (workers and supervisors) goal-setting meetings to set performance-safety goals in regard to housekeeping activities, access to heights, and scaffolding construction. This was followed by publicly displayed feedback charts for eight weeks, based on observations conducted by trained observers. In that period, the weekly-adjusted gap between current levels and designated goals had to be minimised, providing the necessary incentive for change. A meta-analysis by Krispin and Hantula (1996) revealed a strong effect size for behavioural safety interventions, exceeding all other intervention approaches, including training programmes, rule-enforcement, and safety campaigns (McAfee and Winn, 1989; O'Hara et al., 1985). The reported effect size of a 17% increase in safe behaviour resembles meta-analytic results for other OBM interventions having to do with facets such as product quality, customer service and production efficiency (Stajkovic and Luthans, 1997). As noted by the latter authors, the reported effect size indicates substantial behavioural change, exceeding non-behavioural interventions. Jointly, these data indicate that frequently recurring, short-term, and personally relevant rewards can harness known behavioural tendencies in the service of under-performed behaviours by revising its perceived value functions (i.e. safety behaviour is no different from other under-performed role behaviours).

3.1 Modified value functions must be maintained

Behaviour-based safety does not incorporate mechanisms for maintaining modified behaviour after the conclusion of intervention, apart from perpetuating the intervention protocol *ad infinitum*. If such mechanisms are not put in place, the initial value functions will recur after intervention is terminated, resulting in an extinction curve that will reach its asymptote over a period of several weeks (Krispin and Hantula, 1996). By that time, immediate costs of safety behaviour (slower speed, greater discomfort, higher effort) will override any long-term benefits due to the aforementioned behavioural biases. For this reason, most behavioural safety programmes incorporate an open-ended protocol in which rotating cohorts of workers are trained as observers and feedback providers to co-workers (Daniels, 1989; Geller, 1996, 1999; Krause, 1997). Proponents of this approach suggest that, in addition to combating extinction due to permanently modified cost functions, participation in such programmes also increases workers ownership of safety issues (Geller et al., 1996), also known as 'safety citizenship' (Hofmann et al., 2003).

Another approach for coping with post-intervention extinction is supervision-based safety (SBS) intervention (Zohar, 2002; Zohar and Luria, 2003). This intervention is based on the idea that effective line supervisors provide reinforcement contingencies resembling those in OBM and behavioural safety programmes, i.e. they continually specify performance goals, and provide frequent and timely feedback concerning goal attainment (Komaki, 1998; Locke and Latham, 1990). This feedback acquires its rewarding or punitive quality because it predicts personally significant organisational outcomes (i.e. outcomes such as promotion and raises are usually preceded by recurring supervisory approvals while lack of promotion or dismissal are preceded by negative feedback (Bandura, 1986; Stajkovic and Luthans, 1997, 2001). Hence, whereas conventional behaviour-modification interventions depend on external observers and other appointed officials (including co-workers) to provide feedback and incentives, effective supervisors obtain the same information and provide equally frequent incentives as part of their regular routine. Although external intervention can continue indefinitely, it is probably more susceptible to interruption than effective supervisory practices.

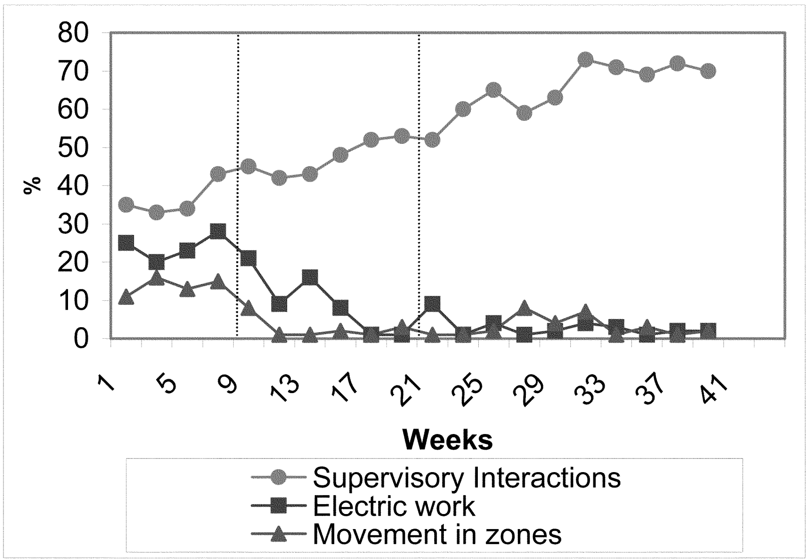
Based on these premises, Zohar and colleagues developed an intervention programme that creates improved supervisory safety practices as leverage for modifying workers' safety behaviours (Zohar, 2002; Zohar and Luria, 2003). Intervention consisted mainly of providing supervisors with individual feedback concerning the frequency of their safety-orientated interactions with subordinates, using an integrated feedback and goal-setting protocol for increasing such supervisory practice. Results indicated a steady, synchronised increase in frequency of both supervisory safety interactions and workers' safety behaviour during intervention, with *continued improvement* during the follow-up months (i.e. after the end of intervention) until it reached a plateau without apparent signs of extinction. Although an experimental design was not used in these field studies, consistent co-variation of supervisory safety interactions and workers' safety behaviour (measured independently by random sampling of supervisory and worker behaviours) supports the notion that modified worker behaviour is sustained by the new supervisory practices. In other words, since intervention had taken place only at the supervisory level, observed changes in workers' behaviour are attributable to a modified value function for safety behaviour stemming from frequent and timely supervisory referrals. When new supervisory safety practices become normative, reflecting supervisory role-redefinition that includes safety as a key performance indicator, the organisation will have established a modified payoff for workers' safety behaviour that counteracts the above-mentioned behaviour biases. Figure 1 provides an example of SBS intervention results from one company.

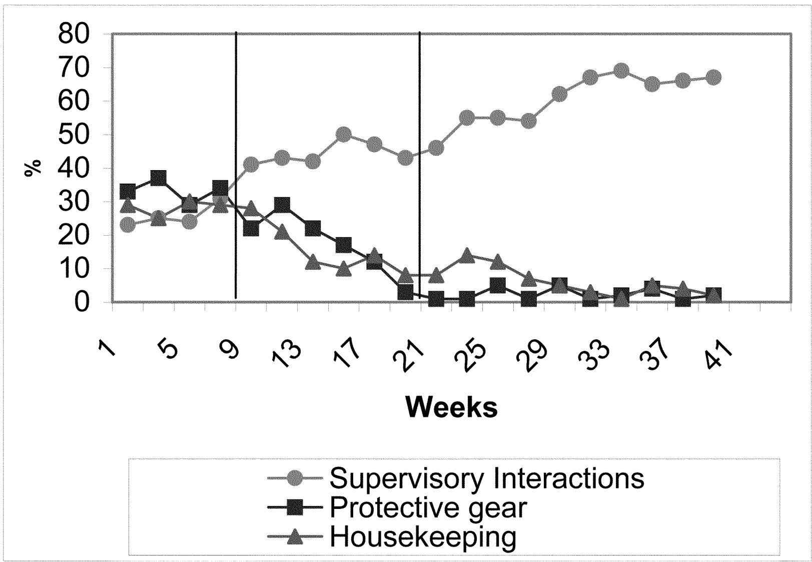
3.2 Safety climate informs behavioural choices

Another area of research supporting our decision-making analysis of safety behaviour derives from the safety-climate literature. Safety climate arises from consensual perceptions of employees concerning the priority of safety in a company or workgroup (i.e. managerial or supervisory commitments), using cumulative experience of employees as the relevant database. Consensus among employees indicating strong managerial emphasis on safety implies a positive climate. A defining characteristic of climate is that it informs role behaviour (i.e. 'which behaviour is likely to be supported and rewarded around here?'). In other words, climate perceptions converge on the priority of key facets such as safety, speed, or quality, which may differ from formal declarations concerning the same issues (Argyris and Schon, 1996). For example, if supervisors repeatedly make safety procedures contingent on production pressures, workers will infer low safety priority even if management's overt policy is that safety has top priority. Similarly, if merit bonuses are awarded to employees who are not known for their safety, other employees will infer low safety priority despite formal declarations to the contrary.

This brief overview suggests an expected-utility link between climate perceptions and role behaviour, stemming from the fact that supervisory practice (which is the source of climate perceptions) is motivationally-relevant because it signals and predicts personally significant organisational outcomes (i.e. recurring supervisory approvals predict longer-term outcomes such as promotion or better job assignments and vice versa (Stajkovic and Luthans, 1997, 2001). Because supervisory action in regard to high-priority issues is relatively frequent and immediate, its collective assessment in the form of safety-climate perceptions should counteract the behavioural biases identified above. Climate perceptions can thus modify the perceived cost function of safety behaviour. This theoretical standpoint is supported by consistent reports of strong climate-behaviour/injury relationships across companies, industrial sectors, and national cultures (Barling et al., 2002; Hofmann and Stetzer, 1996; Neal et al., 2000; O'Dea and Flin, 2001; Zohar, 2000, 2002).

Figure 1 Co-variation of supervisory safety interactions (%) and workers unsafe behavior (%) in two sections of an oil refinery: (a) refinery section; (b) canning and distribution





More direct evidence comes from an intervention study employing the SBS framework described above (Zohar, 2002). Intervention took place in half of the workgroups in a manufacturing company, with the other half serving as control group. Safety climate was measured one month before and one month after a three-month intervention. Whereas climate levels were identical at the beginning of the study, they improved significantly in the experimental group following intervention, while remaining unchanged in the control group. Behaviour-dependent minor injuries, serving as a proxy for safety behaviour, revealed similar changes over the five-month post-intervention period, supporting the role of climate as a social-cognitive counterweight to behavioural biases concerning safety behaviour. (Note, however, that climate is considered the cognitive counterpart, or mediator variable with regard to the main source of safety-behaviour change, i.e. modified supervisory practice and attendant payoff pattern.)

3.3 Ineffective interventions disregard weight modifications

Our analysis suggests that successful intervention programmes must modify perceived value function for safe behaviour by introducing frequent short-term rewards to counteract the tendency to underweight the long-term benefits of safe conduct. Arguably, the reverse is also true. Intervention programmes that disregard the costs involved in safe vs. unsafe behaviour tend to be only partially effective. A case in point is safety training, arguably the most common and resource-demanding intervention strategy (Colligan and Cohen, 2003).

Industrial training programmes promote the acquisition of job-relevant capabilities by developing better knowledge, skills and abilities (Goldstein, 1986, 1989). Knowledge typically refers to theoretical and/or procedural information, e.g. physics principles specifying that friction generates heat, the friction points on a machine or alternative actions for excessive heat. Skill refers largely to psychomotor activities on the job, such as mixing raw materials according to (changing) product specifications, or simultaneously adjusting parameters such as pressure and temperature in combustion chambers. Ability refers mainly to cognitive activities on the job, such as the ability to shift priorities in response to changing conditions, or decision making in situations not covered by existing procedures. The standard training model comprises three stages: assessment of learning needs, provision of training (whose contents and methods are adjusted to initial assessments), and programme evaluation (Goldstein, 1986). Recent expositions of *safety* training methodology follow the standard training model (e.g. (Hilyer et al., 2000; Martin, 1999; ReVelle and Stephenson, 1995)).

The training model focuses on capabilities, even though human performance is a joint function of capabilities *and* motivation (Adams, 1989). If training alone is used as intervention, which is often the case (Salas et al., 1999), this would leave the above-mentioned behavioural tendencies unaffected. Consequently, we should expect limited behavioural change despite increased knowledge or awareness. A meta-analysis of safety interventions supports this line of reasoning, indicating that training (alone), hazard awareness, and safety campaigns only result in limited behaviour change unless combined with a behavioural safety intervention (Krispin and Hantula, 1996). In fact, effect size almost doubles when training is combined with behavioural safety intervention. Similar results were reported by Tracey et al. (1995), indicating that significantly greater role-behaviour change followed training in organisations whose culture encourages continuous learning and personal development. Jointly, these data suggest that if personal costs associated with safe behaviour are not counterbalanced by a contingent reward on the job, new capabilities will not be implemented, regardless of knowledge level by the end of training. This conclusion is corroborated in the transfer-of-training literature, indicating that behaviour change depends largely on whether the (social and physical) work environment encourages, discourages, or actually prohibits the application of new capabilities when trainees get back on the job (Baldwin and Ford, 1988; Tannenbaum and Yukl, 1992; Tracey et al., 1995). As for the prevalence of the training-alone approach, apparently this oversight has been sustained by the fact that most industrial training courses are evaluated through participants' immediate reactions and through objective learning tests collected during training or immediately thereafter, rather than by assessing actual role-behaviour change (Alliger and Janak, 1989; Salas et al., 1999).

Delayed risks require special mention when assessing the efficacy of training as main or sole intervention mode because it has been repeatedly shown that merely informing workers about serious, yet delayed, hazards rarely results in behaviour change even with the best of instruction programmes. A case in point is the use of ear protectors to avert prospective hearing loss some 10–15 years later. A recent literature review indicates that instruction-based training has only limited and largely temporary effect, judged by changes in earplug use among the workforce (Colligan and Cohen, 2003). However, when such instruction is coupled with behavioural safety intervention (Zohar et al., 1980), or supervision-based safety intervention (Zohar, 2002), this will result in a significant improvement in earplug use. Similar results attesting to the inefficacy of instruction-alone protocols were reported with regard to low-back injuries, one of the leading causes of lost-days injuries (Daltroy et al., 1997). In these cases, the extended time-span before symptom appearance results in underweighting safe behaviour in favour of a risky alternative that offers immediate and frequent gains. This attests to the strength of behavioural biases as compared to the utility-maximising alternatives. All of which suggests that risk-perception and behavioural-choice models must incorporate and account for the effect of behavioural biases rather than assuming that self-preservation will override all other considerations.

References

Adams, J.A. (1989) *Human Factors Engineering*, New York: Macmillan.

Alliger, G.M. and Janak, E.A. (1989) 'Kirkpatrick's levels of training criteria: thirty years later', *Personnel Psychology*, Vol. 42, No. 2, pp.331–342.

Argyris, C. and Schon, D.A. (1996) *Organisational Learning II: Theory, Method, and Practice*, Reading, MA: Addison-Wesley.

Baldwin, T.T. and Ford, J.K. (1988) 'Transfer of training: a review and directions for future research', *Personnel Psychology*, Vol. 41, No. 1, pp.63–105.

Bandura, A. (1969) *Principles of Behavior Modification*, New York: Holt, Rinehart & Winston.

Bandura, A. (1986) *Social Foundations of Thought and Action*, Englewood Cliffs, NJ: Prentice-Hall.

Barkan, R., Zohar, D. and Erev, I. (1998) 'Accidents and decision making under uncertainty: a comparison of four models', *Organisational Behavior and Human Decision Processes*, Vol. 74, No. 2, pp.118–144.

Barling, J., Loughlin, C. and Kelloway, E.K. (2002) 'Development and test of a model linking transformational leadership and occupational safety' *Journal of Applied Psychology*, Vol. 87, No. 3, pp.488–496.

Barron, G. and Erev, I. (2003) 'Small feedback-based decisions and their limited correspondence to description-based decisions', *Journal of Behavioral Decision Making*, Vol. 16, No. 3, pp.215–233.

Barron, G., Yechiam, E. and Erev, I. (2004) *The judgment-decision paradox and the contingent recency effect*, Technion Working paper No. 04–05.

Bornstein G. and Ben-Yosseff, M. (1994) 'Cooperation in inter-group and single group prisoner dilemmas', *Journal of Experimental Social Psychology*, Vol. 30, No. 1, pp.52–67.

Colligan, M.J. and Cohen, A. (2003) 'The role of training in promoting workplace safety and health', in J. Barling and M. Frone (Eds), *The Psychology of Workplace Safety* (pp.223–248). Washington, DC: American Psychological Association.

Daltroy, L.H., Iversen, M.D. and Larson, M.G. (1997) 'A controlled trial of an educational program to prevent low-back injuries', *New England Journal of Medicine*, Vol. 337, No. 5, pp.322–328.

Daniels, AC. (1989) *Performance Management*, Tucker, GA: Performance Management Publications.

Dawes R.M. (1980) 'Social dilemma', *Annual Review of Psychology*, Vol. 31, pp.169–193.

Erev, I. and Roth, A.E. (1998) 'Predicting how people play games: reinforcement learning in games with unique strategy equilibrium', *American Economic Review*, Vol. 88, No. 4, pp.848–881.

Erev, I., Bornstein, G. and Galili, R. (1993) 'Constructive intergroup competition as a solution to the free-rider problem in the workplace', *Journal of Experimental Social Psychology*, Vol. 29, No. 6, pp.463–478.

Erev, I., Gopher, D., Itkin, R. and Greenshpan, Y. (1995) 'Toward a generalization of Signal Detection Theory to n-person games: the example of two person safety problem', *Journal of Mathematical Psychology*, Vol. 39, No. 4, pp.360–376.

Fischhoff, B., Parker, A.M., Bruine De Bruin, W., Downs, J., Palmgren, C., Dawes, R. and Manski, C.F. (2000) 'Teen expectations for significant life events', *Public Opinion Quarterly*, Vol. 64, No. 2, pp.189–205.

Frederick, S., Loewenstein G. and O'Donoghue T. (2002) 'Time discounting and time preference: a critical review', *Journal of Economic Literature*, Vol. 40, No. 2, pp.351–401.

Geller, E.S. (1996) *The Psychology of Safety: How to Improve Behaviors and Attitudes on the Job*, Randor, PA: Chilton Books.

Geller, E.S. (1999) 'Behavior-based safety: confusion, controversy, and clarification', *Occupational Health & Safety*, Vol. 68, No. 1, pp.40–50.

Geller, E.S., Roberts, D.S. and Gilmore, M.R. (1996) 'Predicting propensity to actively care for occupational safety', *Journal of Safety Research*, Vol. 27, No. 1, pp.1–8.

Goldstein, I.L. (1986) *Training in Organisations* (2nd Ed), Monterey, CA: Brooks-Cole.

Goldstein, I.L. (1989) *Training and Development in Organisations*, San Francisco, CA: Jossey-Bass.

Greening, L. (1997) 'Risk perception following exposure to a job-related electrocution accident', *Acta Psychologica*, Vol. 95, No. 3, pp.267–277.

Harpaz, I. and Snir, R. (2003) 'Workaholism: its definition and nature', *Human Relations*, Vol. 56, No. 3, pp.291–319.

Heinrich, H.W. (1931) *Industrial Accident Prevention: a Scientific Approach*, New York: McGraw-Hill.

Herrnstein, R.J., Loewnstein, G.F., Prelec, D. and Vaughan, W. (1993) 'Utility maximization and melioration: internalities in individual choice', *Journal of Behavior & Decision Making*, Vol. 6, No. 3, pp.149–185.

Hertwig, R. Barron, G., Weber, E. and Erev, I. (2004) 'Decisions from experience and the effects of rare events in risky choice', *Psychological Science*, Vol. 15, No. 8, pp.534–539.

Hilyer, B., Veasey, D., Oldfield, K. and McCormick, L. (2000) *Effective Safety and Health Training*, Boca Raton, FL: CRC Press.

Hofmann, D.A. and Stetzer, A. (1996) 'A cross-level investigation of factors influencing unsafe behaviors and accidents', *Personnel Psychology*, Vol. 49, No. 2, pp.307–339.

Hofmann, D.A., Morgeson, F.P. and Gerras, S.J. (2003) 'Climate as a moderator of the relationship between LMX and content specific citizenship: safety climate as an exemplar', *Journal of Applied Psychology*, Vol. 88, No. 1, pp.170–178.

Kahneman, D. and Tversky, A. (1979) 'Prospect theory: an analysis of decision under risk', *Econometrica*, Vol. 47, No. 2, pp.263–291.

Komaki, J.L. (1998) *Leadership from an Operant Perspective*, New York: Routledge.

Krause, T.R. (1997) *The behavior-Based Safety Process* (2nd edition). Ojai, CA: BST.

Krispin, J. and Hantula, D.A. (1996) 'A meta-analysis of behavioral safety interventions in organisations', Proceedings of the 1996 annual meeting of the Eastern Academy of Management, Philadelphia, PA.

Lichtenstein, S., Slovic, P., Fischhoff, B. and Combs, B. (1978) 'Judged frequency of lethal events', *Journal of Experimental Psychology: Human Learning and Memory*, Vol. 46, No. 4, pp.551–578.

Lingard, H. and Rowlinson, S. (1997) 'Behavior-based safety management in Hong Kong's construction industry', *Journal of Safety Research*, Vol. 28, No. 4, pp.243–256.

Locke, E.A. and Latham, G.P. (1990) *A Theory of Goal Setting and Task Performance*, Englewood Cliffs, NJ: Prentice-hall.

Luthans, F. and Kreitner, R. (1985) *Organisational Behavior Modification and Beyond*, Glenview, IL: Scott, Foresman.

Martin, K.P. (1999) 'Health and safety instruction', in L.J. DiBerardinis (Ed.), *Handbook of Occupational Safety and Health* (2nd edition), pp.517–537). New York: Wiley.

Maslow. A. (1970) *Motivation and Personality* (2nd Ed.) New York: Harper and Row.

McAfee, R.B. and Winn, A.R. (1989) 'The use of incentives/feedback to enhance workplace safety: a critique of the literature', *Journal of Safety Research*, Vol. 20, No. 1, pp.7–19.

Messick, D.M. and Brewer, M.B. (1983) 'Solving social dilemmas: a review', in L. Wheeler and P. Shaver (Eds), *Review of Personality and Social Psychology*, Vol. 4, pp.11–44, Beverly Hills, CA: Sage.

National Safety Council (1999) *Injury Facts*, Itasca, IL: National Safety Council.

Neal, A., Griffin, M.A. and Hart, P.M. (2000) 'The impact of organisational climate on safety climate and individual behavior', *Safety Science*, Vol. 34, No. 1, pp.99–110.

O'Dea, A. and Flin, R.H. (2001) 'Site managers and safety leadership in the oil and gas industry', *Safety Science*, Vol. 37, No. 1, pp.39–57.

O'Hara, K., Johnson, C.M. and Beehr, T.A. (1985) 'Organisational behavior management in the private sector: a review of empirical research', *Academy of Management Review*, Vol. 10, No. 4, pp.848–864.

Perry, O., Erev, I. and Haruvy, E. (2002) 'Frequent delayed probabilistic punishment in law enforcement', *Economics of Governance*, Vol. 3, No. 1, pp.71–85.

Rapoport, A. and Moshowich, L. (1966) 'Experimental studies of stochastic models for the prisoner dilemma', *Behavioral Science*, Vol. 11, No. 6, pp.444–458.

Rapoport, A. and Chammah, A.M. (1965) *Prisoner's Dilemma: A Study in Conflict and Cooperation*, Ann Arbor: University of Michigan Press.

Reason, J. (1997) *Managing the Risks of Organisational Accidents*, Aldershot, UK: Ashgate.

ReVelle, J.B. and Stephenson, J. (1995) *Safety Training Methods* (2nd edition), New York: Wiley.

Salas, E., Cannon-Bowers, J., Rhodenizer, L. and Bowers, C. (1999) 'Training in organisations: myths, misconceptions, and mistaken assumptions', in G.R. Ferris (Ed.), *Research in Personnel and Human Resources Management*, Vol. 17, pp.123–161.

Skinner, B.F. (1974) *About Behaviorism*, New York, NY: Vintage.

Stajkovic, A.D. and Luthans, F. (1997) 'A meta-analysis of the effects of organisational behavior modification on task performance', 1975–95, *Academy of Management Journal*, Vol. 40, No. 5, pp.1122–1149.

Stajkovic, A.D. and Luthans, F. (2001) 'Differential effects of incentive motivators on work performance', Academy of Management Journal, 44 (3), 580–590.

Tannenbaum, S.I. and Yukl, G.A. (1992) 'Training and development in work organisations', *Annual Review of Psychology*, Vol. 43, pp.399–441.

Thaler, R.H. and Shefrin, H.M. (1981) 'An economic theory of self control', *Journal of Political Economics*, Vol. 89, No. 2, pp.392–406.

Thorndike, E.L. (1898) 'Animal intelligence: an experimental study of the associative processes in animals', *Psychological Monographs*, Vol. 2.

Tracey, J.B., Tannenbaum, S.I. and Kavanagh, M.J. (1995) 'Applying trained skills on the job: the importance of the work environment', *Journal of Applied Psychology*, Vol. 80, No. 2, pp.239–252.

Viscusi, K.W. (1992) *Smoking: Making the Risky Decision*, New York: Oxford University Press.

Yechiam, E., Erev, I. and Haruvy, E. (2002) 'Reinforcement learning and the prevention of data catastrophes', *Journal of Managerial Psychology*, Vol. 17, No. 7, pp.599–611.

Zohar, D. (2000) 'A group-level model of safety climate: testing the effect of group climate on micro-accidents in manufacturing jobs', *Journal of Applied Psychology*, Vol. 85, No. 4, pp.587–596.

Zohar, D. (2002) 'Modifying supervisory practices to improve sub-unit safety: a leadership-based intervention model', *Journal of Applied Psychology*, Vol. 87, No. 1, pp.156–163.

Zohar, D. and Luria, G. (2003) 'The use of supervisory practices as leverage to improve safety behavior: a cross-level intervention model', *Journal of Safety Research*, Vol. 34, No. 5, pp.567–577.

Zohar, D. and Luria, G. (2005) 'A multilevel model of safety climate: Cross-level relationships between organization and group-level climates', *Journal of Applied Psychology*, Vol. 90, pp.616–628.

Zohar, D., Cohen, A. and Azar, N. (1980) 'Promoting increased use of ear protectors in noise through information feedback', *Human Factors*, Vol. 22, No. 1, pp.69–79.

Notes

[1] The tendency to *underweight* rare events is limited to recursive decisions that are based on cumulative experience of the type discussed above. *Overweighing* of rare events occurs when a single choice must be made based on a description of possible alternatives (e.g. Kahneman and Tversky, 1979, studies in conjunction with Prospect Theory). Thus, 83% of participants in Kahneman and Tversky's study prefer the prospect of 'a sure loss of 5' over a gamble described as 'a loss of 5000 with probablilty1/1000; no loss otherwise'.

[2] Other popular names to these social problems include prisoner dilemmas and public-good problems.

[3] An exception to this pattern occurs in a two-person prisoner dilemma game in which two players interact repeatedly over many trials. In this case reciprocation may emerge (Rapoport and Chammah, 1965; Rapoport and Moshowich, 1966).