1-1-2010

On Effect Assessment in Work Environment Interventions – A Literature Overview and Methodological Reflection

W. Patrick Neumann
Ryerson University, Canada, pneumann@ryerson.ca

Jörgen Eklund
Royal Institute of Technology, Sweden, Jorgen.Eklund@sth.kth.se

Bo Hansson
Mälardalen University, Sweden, Bo.Hansson@oecd.org

Lars Lindbeck
Royal Institute of Technology, Sweden, lars.lindbeck@ownit.nu

Follow this and additional works at: http://digitalcommons.ryerson.ca/ie

Part of the Community Health and Preventive Medicine Commons, Ergonomics Commons, Investigative Techniques Commons, and the Occupational Health and Industrial Hygiene Commons

Recommended Citation
http://digitalcommons.ryerson.ca/ie/2

This Article is brought to you for free and open access by the Industrial Engineering at Digital Commons @ Ryerson. It has been accepted for inclusion in Industrial Engineering Publications and Research by an authorized administrator of Digital Commons @ Ryerson. For more information, please contact bcameron@ryerson.ca.
On Effect Assessment in Work Environment Interventions – A Literature Overview and Methodological Reflection

Neumann, W.P.¹, Eklund, J.²,₅, Hansson, B.³, Lindbeck, L.⁴

1- Department of Mechanical and Industrial Engineering, Ryerson University, Canada
2- Department of Management and Engineering, Linköping University, Sweden
3- School of Business, Mälardalen University, Västerås, Sweden
4- Verkstadsgatan 10, 117 36 Stockholm, Sweden
5- Royal Institute of Technology, KTH Syd, SE-141 52 Huddinge, Sweden

Corresponding Author:
Dr. W.P. Neumann, Eur.Erg.
Mechanical and Industrial Engineering
Ryerson University
350 Victoria St., Toronto, ON,
Canada, M5B 2K3

Tel.: +1 – 416 – 979 5000 x7738
Fax: +1 – 416 – 979 5265
pneumann@ryerson.ca

An earlier report on this project, in Swedish, is available:


A subsequent version of this work is also available:

ABSTRACT
Despite many positive case studies our overview of the work environment intervention (WEI) research literature finds mixed results. There is support for the profitability of WEI investments such as training and personnel policies at the organisational level but less clear results for disorder reductions. The financial benefits of WEIs were greater for performance gains than for reduced sickness costs. Multifactor interventions are widely seen as key to successful intervention but are difficult to evaluate and unused in experimental studies. Review inclusion criteria excluding studies with good interventions but non-experimental evaluations, pose a quality criteria selection bias. Difficulties in proving WEI effectiveness may depend on views of what constitutes good scientific quality. WEI effects are clear in some cases but are difficult to show in others. Evaluation poses methodological challenges that contribute to the lack of clear evidence for WEI effectiveness. There is a need for more practical multifactor WEIs and non-experimental evaluation strategies suited to today’s complex systems.

RELEVANCE:
Ergonomists and managers should understand that the problems in 'proving' the effectiveness of ergonomics are related to perceptions of what constitutes proof. Progress in the practise of ergonomics should recognise the difficulty of organisational change, the weaknesses of experimental traditions, and the need for multifactor interventions that reach deep into the work process to maximise impact. Isolating effects is difficult but this does not mean no effects exist.

Key words: Intervention, health promotion, review methodology, ergonomics research, musculoskeletal disorders, economics
1.0 Introduction

An increasing number of literature reviews are being published in scientific journals. In a comparison between reviews, there are sometimes substantial differences in the conclusions drawn. In the field of ergonomics, this has been apparent (see Cole et al., 2005; Hansson and Westerholm, 2001; Hignett, 2003; Karsh et al., 2001; Kuorinka and Forcier, 1995; Westgaard and Winkel, 1997). Many of the reasons behind these differences seem to be how the original papers have been evaluated, and thus included or excluded in the review. The different conclusions are problematic for practitioners who wish to apply scientific findings, and this is therefore an issue with high priority.

Evaluation is a diverse concept. Evaluations may only fulfill a symbolic need (Vedung, 1998), they may be quantitative or qualitative (Patton, 2002), and summative or formative (Rossi et al., 1999). Effects might be calculated, or for more credibility, measured. Effects might also be intended or unintended, there can be side-effects that are outside the aims of the intervention, perverse effects that go against the aims, and might be positive or negative in nature (Vedung, 1998). Evaluation is in itself, a broad field with its own research journals (Evaluation, Evaluation Review, American Journal of Evaluation, Evaluation and Program Planning, Evaluation Practice etc.) and traditions dealing with the challenges of proving that an intervention causes a particular effect.

One of the strongest approaches for demonstrating causality comes from the medical experimental tradition involving the use of randomisation and control groups. Randomised control trials (RCT), seen as the strongest study design, also suffer from many drawbacks including: Vulnerability to differential changes (e.g. drop out), context specificity (they are not generaliseable), experimental programs differ from ‘real’ program (not tranferrable), RCTs do not apply to full coverage programs, they have high cost, spill-over in the workplace can cause contamination and, finally, RCTs are not suitable in early stages of implementation (Rossi et al., 1999). In many workplaces RCT are not realistic and alternative designs such as before-after trials without or (better) with
control groups are more appropriate. Longitudinal case studies that track variables over longer time periods can also provide useful information on cause-effect relationships.

Text Box 1: Hill’s (1965) viewpoints on indicators for assessing causality in an individual disease model.

Hill (1965) has outlined criteria for determining the plausibility of causality relationship in terms of the sources of disease for individuals (see Text Box 1). Interventions at the organisational or societal level represent the most complex types of interventions that can have many interacting elements and synergistic interactions whose effects are difficult, if not impossible, to isolate (Ekberg, 1994). Such interventions operate at a level of system complexity one order higher than the individual humans whose health is to be improved (Skyttner, 2001). The more complex the system subjected to intervention is, the more complex the intervention itself needs to be, and the more difficult the evaluation will be. Sometimes it is not possible to conduct an experimental evaluation. Furthermore complexity in systems implies that relationships will vary widely between systems (in this case organisations) and that these relationships are unstable and may shift over time (Skyttner, 2001; Backström et al., 2002). Under these circumstances the usefulness positivistic concept of ‘proof of effectiveness’ can be questioned since generaliseability is not possible – there is no single ‘general’ case on which to intervene (Gustavsen et al., 1996). It is appropriate therefore to question the purpose of the evaluation and consider a shift in focus from ‘proof of effectiveness’ to the generation of information that can guide further action (Patton, 1997). Those outside the intervention must consider the transferability of findings to their own context, while those inside the intervention need to ensure that the efforts to improve working environment are adapted to the changes

1. Strength of association
2. Consistency of association
3. Specificity or association
4. Time sequence of association
5. Dose-response association
6. Biological plausibility
7. Coherence with theory
8. Experimental results
9. Other experience
ongoing in the organisation. This research approach is more consistent with research in the areas of business practices, such as Total Quality Management where the emphasis is less on ‘proving effectiveness’ and more on determining what features of implementation appear to be associated with better performance (Ennals, 1999).

This paper discusses methodological issues emerging from an overview of the ergonomics and working environment intervention (WEI) research literature. This is a condensed and modified version of a larger, Swedish language, report (Eklund et al., 2006) prepared as part of the SMARTA research programme (English expansion: ‘Strategies and Methods for Management of the Working Environment’ (Christmansson et al., 2005)). The term WEI used in this paper is rooted in Scandinavian conceptions of workplace health and safety and is similar to the broad international definition of ‘ergonomics’ as it applies to workplaces (IEA Council, 2000). In this case we also include elements such as workplace training or human resource policies that might not be thought of as ‘ergonomics’ by some readers. We aim to use the literature overview to examine the problems faced in evaluation of the effects of WEIs ergonomics interventions and discuss methodological issues that may be contributing to the diverging conclusions regarding the effectiveness.

2.0 Literature Overview

WEIs generally include a focus on one or more aspects including: the technical system (e.g. equipment), the organisational system (e.g. job rotation), or behaviour related aspects (e.g. training programs) (Karsh et al., 2001). While Swedish law requires a goal of “preventing illness and accidents at work while developing a satisfying working environment” (Arbetsmiljöverket, 2001), there has also been growing interest in the economic effects of WEI (Hendrick, 1996; Riel and Imbeau, 1997; Beevis, 2003; Stanton and Baber, 2003; Oxenburgh et al., 2004). For these reasons we have conducted reviews of available literature on both the wellness and the economic agendas for WEI. The review focussed on literature since 1995 including both individual studies and available review articles. This overview attempts to gain an overview of the topic and thus no
papers were excluded based on study design criteria as might be done in a ‘systematic’ review. We did however include such reviews here (van Poppel et al., 1997) and even considered reviews of reviews (Källenstål et al., 2004).

2.1 Musculoskeletal disorder (MSD) interventions

That workplace factors can cause MSDs has been known for over 300 years (Ramazzini, 1700). Nevertheless the last quarter of the 20th century saw a tremendous amount of research on the physical and psychosocial risk factors for MSDs and a number of excellent reviews exist (Bongers et al., 1993; Bernard, 1997; Buckle and Deveraux, 1999; Ariens et al., 2000; de Beeck and Hermans, 2000; Hoogendoorn et al., 2000b; Netherlands, 2000; Malchaire et al., 2001a; Buckle and Deveraux, 2002). Studies continue to corroborate (consistency; one of Hill’s causality criteria; Text box 1) the relation between workplace demands and MSDs to the back (Hoogendoorn et al., 2000a; Hoogendoorn et al., 2001; Kerr et al., 2001), neck (Ariens et al., 2001a; Ariens et al., 2001b), neck & shoulder (Fredriksson et al., 2000; Östergren et al., 2001); and hand-wrist (Malchaire et al., 2001b). This causal evidence notwithstanding, it is a challenge to ‘prove’ that workplace interventions, run with the intention of reducing these MSD risk factors in the workplace, have led to reduced MSD rates. While older reviews of intervention research found substantial effects (Snook, 1987), newer reviews are more cautious (Tveito et al., 2004; Verhagen et al., 2004). It should be noted that these newer reviews, as we will discuss later, had different inclusion criteria emphasizing experimental methods. We highlight below some of the specific MSD intervention approaches that have been studied and for which a summary statement seems warranted:

**Backbelts** – The use of backbelts has not been seen to be effective in preventing back pain at work (van Poppel et al., 1997; Källenstål et al., 2004; Silverstein and Clark, 2004).

**‘Back-School’ programs** – Were not generally seen as effective in preventing MSDs in industry (van Poppel et al., 1997; Källenstål et al., 2004; Silverstein and Clark, 2004), or in the hospital sector (Hignett, 2003; Hignett et al., 2003).
Physical Training – Was seen to have some benefits in preventing back pain (van Poppel et al., 1997; Källestål et al., 2004; Silverstein and Clark, 2004) and neck pain (Källestål et al., 2004).

Biomechanical Interventions – While reviews of ‘risk factor modification’ attempts, including biomechanical risks, did not find proof of consistent benefits the studies were mostly of insufficient experimental quality to provide such proof. WEI outcome measures are most often in terms of risk factor exposure and it is left to the epidemiological studies to prove the link between the risk factor and MSDs (e.g. Karsh et al., 2001). One study reported that interventions aimed at reducing physical work-load were seen to yield significant reductions in low back disorders in workers engaged in manual handling (Marras et al., 2000) – a reversal (time sequence; Hill’s 4th causality indicator) of the epidemiological linkage that is uncommon.

Multi-component Interventions – Multifactor interventions were seen by reviewers to be more likely to succeed in preventing MSDs than simpler single component interventions (Guastello, 1993; Westgaard and Winkel, 1997; Karsh et al., 2001; Hignett, 2003; Hignett et al., 2003; Silverstein and Clark, 2004).

Study quality – Study quality was generally seen as ‘low’ when compared to medical experimental traditions. Van Poppel et al. found only 11 studies with control groups and none using ‘blinding’ techniques such as those used in pharmacological studies. While systematic reviews have justified the exclusion of many studies due to ‘quality’ deficits, there exist many studies that have indeed shown positive effects of interventions aimed at MSD prevention. Kilroy and Dockrell (Kilroy and Dockrell, 2000) present a typical case study of this kind with pre-post measures, no control group, and follow-up shortly after the intervention. Studies with more experimental designs, using control groups for example, generally showed modest effects (Morken et al., 2002). Laing et al. provide another example of an ambitious intervention with control group that was unable to achieve substantial reductions in perceived effort or pain levels (Laing et al., 2005). Possible explanations for this included the short follow-up time (10 months) and the
ongoing changes in production rates and staffing that was different between the intervention and control groups. Feuerstein et al. (2004) had a different problem – both their intervention group (workplace modifications) and the control group (who received training in stress management) demonstrated significant reductions in reported MSD symptoms even after 12 months. While such changes might be explained as Hawthorne effect, it is also possible that both interventions were effective in this case – another ‘problem’ with the use of control groups in these settings. Volinn points out that the quality of a studies’ evaluation design is often inversely related to the reported effect (Volinn, 1999). Few studies addressed MSD prevention at the organisational level.

2.2 Outcomes with Economic implications

We now turn our attention from the health effects of intervention to the economic effects of intervention. This is consistent with the ‘system performance’ objective in the international definition of ‘ergonomics’ (IEA Council, 2000) and has become an increasingly discussed objective inspiring special journal editions on the topic (Stanton and Baber, 2003).

2.2.1 Health Interventions

On the cost side, studies have shown dramatic increases in health insurance costs – increases of 60% since 2000 have been seen in the US (Baicker and Chandra, 2005). Relatively few studies have been able to study the relation between health interventions and productivity or quality performance. Nevertheless, many overview studies have shown that efforts to reduce sickness absence and improve employee’s health are profitable for the company (see for example: Goetzel et al., 1999; Pelletier, 1999; Aldana, 2001; Anderson et al., 2001; Golaszewski, 2001; Karsh et al., 2001; Riedel et al., 2001). These interventions are typically aimed at the individual level and rarely cover the organisational level. In most instances when the organisational level is considered in an intervention the focus is usually on the physical work environment rather than psychosocial factors (Shain and Kramer, 2004). Individual interventions can include everything from smoking cessation programs, fitness initiatives, or nutrition improvement initiatives to improve employee’s health. While these relatively small
investments can be cost-beneficial even with relatively small effects, most case reports are from project personnel rather than independent researchers and should be interpreted with caution. The application of individual level randomised control studies is probably not appropriate in most cases as the selection of individuals for a health promotion intervention may have an effect on both the control group individuals as well as those receiving intervention thus biasing the study results. Furthermore the transferability of economic benefit findings will depend on the extent to which companies bear the costs for sickness of their employees. In Sweden, for example, where companies bear less sickness absenteeism costs than their American counterparts, the cost-benefit profile of interventions may be considerably different.

2.2.2 Job Satisfaction and good management
Many studies have documented a strong correlation between employee satisfaction and customer satisfaction. Wiley (1991), for example, found that hardware store customers perception of service and intention to shop in that store in the future was strongly related to employees perceptions of the working environment (physical environment, relation to co-workers etc.). Furthermore in pharmacy retail outlets; satisfied employees lead to satisfied customers and also to improved productivity (Westlund and Löthgren, 2001). Bernhardt et al. (2000), in a rare longitudinal analysis, demonstrated that improved customer satisfaction lead to improved profitability in fast food restaurants. A counter-example, demonstrating higher profitability in food retail stores with less satisfied employees (Silvestro, 2002), also exists suggesting there is a complex relationship between these factors.

A related thread of research has found that leadership plays a strong role in job satisfaction and well being (Aronsson et al., 2004) and that a person centred leadership contributes to improved cost-effectiveness and adaptive ability (Arvonen and Pettersson, 2002). In school environments studies have shown productivity to be related to both good leadership (Griffeth et al., 2000) and a good psychosocial environment (Tarter and Hoy, 2004). While these studies are not ‘interventions’ per se, there appear to be useful lessons on the role work environment plays for both individual well being and the
economic performance of the organisation. The empirically observed linkages are consistent with theoretical descriptions of the ‘Service Profit Chain’ (Heskett et al., 1994) and demonstrate how far reaching the benefits of work environment interventions can be. This ‘reach’ of WEI in the organisation complicates evaluating all the benefits of WEIs.

2.2.3 Personnel Training

Studies in the field of labour economics, conducted in different countries using large databases which include data on training investments and economic variables, have generally shown that personnel training has positive economic effects in terms of productivity, profitability, and company market value. Recent empirical evidence on the profitability of training investments is largely consistent across countries and includes studies by Barrett and O’Connell (2001), Dearden et al. (2000), Groot (1999), Bosworth and Loundes (2002), D’Arcimoles (1997), Bassi et al. (2004), and Hansson et al. (2004). Most of these studies are based on longitudinal ‘panel’ datasets and the lagged effects suggest a causal relation (Hill’s criteria #4 Text Box 1) – which goes beyond the simple explanation that more profitable companies can afford to train their personnel. An international study of 26 countries found that 3% of labour costs is invested in employee training and that 45% of all employees receive some kind of training every year (Hansson, 2007). In Sweden the amount invested in employee training was almost 6 times that invested in work environment improvements (Hansson, 2003). Both forms of investment appear to depend on economic factors and can drop drastically (40-50%) in times of financial difficulty.

2.2.4 High Performance Work Systems (HPWS)

‘High performance work systems’ has emerged from the mid 1990’s in the Human Resource Management (HRM) field and there are hundreds of studies on the topic. HPWS is largely focussed on employee selection methods, pay- and reward systems, training of employees, and how the company involves and motivates its employees. The theoretical basis of HPWS is that the firms profitability is influenced by how well these systems are developed to align with overall company strategies. For seminal papers in the area see Huselid (1995), Becker and Gerhart (1996), Becker and Huselid (1997) or
Delaney and Huselid (1996). HPWS systems are sometimes referred to as ‘high commitment policies” or “Human capital enhancing systems” and carry parallels to involvement approaches described as ‘participatory ergonomics’ (Noro and Imada, 1991; Haimes and Carayon, 1998). Despite variations in emphasis for HPWS there are many studies demonstrating significant positive relationships between a well developed human resource systems and firm profitability (Boselie et al., 2001). Methodologically the research studies are mostly cross-sectional in nature and lack the robust longitudinal databases used to examine employee training effects (previous section). Exceptions include the studies if Ichniowski et al. (1997) and d’Arcimoles (1997). The reliance on cross sectional designs is in part due to the nature of HRM systems - since companies rarely change their HRM system, or change them very slowly, there is little opportunity to study changes over time with respect to changes in firm profitability. In the study of Ichniowski et al. (1997), longitudinal data was collected from 26 American steel manufacturers encompassing 36 production lines. Lines were classified as being either ‘innovative’ or ‘conservative’ in their HRM practices on a 1-4 scale. Innovative lines used incentive pay schemes, team-work, careful screening of new employees, employment security, job flexibility, training, and information sharing while ‘conservative’ lines were characterised by more traditional approaches with narrow job definitions, strict work rules, and hourly pay with close supervision. Results showed that ‘innovative’ lines had 7% more production time (less downtime) and higher quality than did ‘conservative’ lines – where each percentage point increases monthly income $30,000 and costs $2100 ($27,900 monthly profit increase). Furthermore, those companies that moved between categories (as a result of policy changes) showed significant effects similar to those of cross sectional studies but with the advantage of observing the time sequence – improving working environment (towards ‘innovative’ lines) led to improved performance.

2.2.5 Work environment interventions
A Number of studies have demonstrated linkages between efforts to improve the work environment and productivity. Anderzén & Arnetz (2005), for example, describe an intervention in tax-office workers that included a broad range of work organisational
factors and found improvements in perceived stress, biological stress markers, sickness absences and productivity. Unfortunately productivity could only be measured in 12 of 22 sites and was not significant due to few observations. In a sub-sample of 1139 companies from the massive (25,000 projects) Swedish ‘Working Life’ fund Gustavsen at al. (Gustavsen et al., 1996) found that those companies that had more substantial work organisational improvements (such as increased operator participation) also showed higher increases in productivity. Johanson (1997), in another review of the ‘Working Life’ fund projects, found that reductions in short and long term sickness absence along with productivity improvements among randomly selected companies yielded payback periods of 2.0 years for private sector and 4.8 years for public sector organisations. Productivity gains made the largest contribution to the economic effects. While these results depended on company self reports the self reports for sickness absence agreed well with National statistics from this period. Similarly, Abrahamsson (2000) in a well conducted multifactorial intervention in the steel industry showed a 2.2 year payback period (59% quality savings, 39% productivity improvement, and 2% sickness absence cost).

Many case studies have also reported that the largest economic effects from work environment interventions can be found in productivity and quality improvements (Helander and Burri, 1995; Hendrick, 2003; Oxenburgh et al., 2004). These cases however are often based on calculated rather than actual results, rarely include data from reference (control) groups, and frequently lack detail in description of methodologies. Nevertheless a number of studies have demonstrated the economic benefits of WEI (see for instance: Helander and Burri, 1995; Kemmlert, 1996; Johanson, 1997; Kompier et al., 1998; Drury, 1999; Arnetz et al., 2003; Beevis, 2003; Butler, 2003; Sen and Yeow, 2003; Lahiri et al., 2005).

3.0 Discussion
How can one measure, identify, or evaluate the effects of efforts to improve the working environment (WE)? While this will depend on context, a better result is likely if evaluation is planned at the same time as the intervention (Vedung 1998). Validity of the
evaluation will be improved if each stage along the hypothesized chain of effects is considered. For example:

1. Measure or describe the WE itself
2. Measure or describe the individuals immediate experience of the WE
3. Measure or describe individuals reactions to the WE (e.g. turnover, job changes)
4. Measure or describe health consequences for the individual (long-term)
5. Measure or describe system effects in terms of safety, productivity or quality

Each level provides different insight along the hypothesised effect pathway and can be considered as providing indicators at different time points – leading indicators first with lagging indicators later in the causal chain (Cole et al., 2003). Any evidence from along the chain can provide insight into the effects of any change efforts. Karsh (2001) has argued that, since many high quality studies have demonstrated the link between risk factors and MSDs, then it is sufficient in an intervention to demonstrate a reduction in risk factor exposures. In this position the emphasis is no longer on reinforcing the epidemiological evidence but on intervention itself.

The etiologically preventable fraction of ill health also poses a challenge for using health outcomes to evaluate interventions. In epidemiological studies it is unusual for any 1 exposure variable to account for more than 10% of the injury variance. If the intervention only partially eliminates this factor then the possible impact that can be expected is quite small. The research implication is that the sample sizes needed to spot this difference become enormous. This problem has also been framed in terms of the ‘intensity’ of the intervention – the extent to which the intervention makes real and substantial change in risk factor exposure amongst the workforce population being served (Cole et al., 2003).

*Time an Issue* - Time becomes a critical issue in evaluating WE interventions. Cancer for example can have a 30 year latency. MSDs can also have a long ‘incubation’ period and it can also take a long time for disorders to resolve. Some MSDs must be considered as chronic and won’t be resolved by preventive measures. If permanent damage has
occurred then the best an intervention can hope for is to prevent the injury from worsening – even as the existing MSD remains. While the elimination of certain risk factors may be achieved and measured, it is more difficult to demonstrate the effects of these changes on musculoskeletal symptoms or complaints. To measure the prevalence and intensity of MSDs is itself difficult. To apply these measures as indicators of the effects of an intervention is even more problematic. Furthermore, some interventions can take a long time before they are fully integrated in the company and start to have an effect on the work system – 3 to 5 years has been suggested for organisational change efforts (Toulmin and Gustavsen, 1996). Thus long evaluation periods are needed to identify effects; although within this period other interventions in the company or economy might overshadow the effects of a WE intervention. The need for long study periods may conflict with research funding agencies that have shorter funding windows.

From a systems theory perspective, both the initial conditions of a complex system and the ongoing changes within that system over time can have pronounced effects on how a given intervention may affect that system (Skyttner, 2001; Backström et al., 2002). Empirical reports seem consistent with this view suggesting that intervention effects can be modified or compromised by macro-economic changes, management culture, and a companies current rationalisation efforts (Bao et al., 1996; Polanyi et al., 2005). Interventions can also be influenced by the more micro issues of normal life events such as marriage, parenting, and the death of personnel engaged in the change process (Neumann, 2004). Relationships between system elements also change over time – strength and spinal load-tolerance, for example, change with age and fitness level. Another aspect of time is that many core features of the company, and hence the working environment, change very slowly or only rarely (such as the HR systems). These changes are often not clearly delineated in time but form an ongoing developmental process in the company. This makes the study of these rare and slow changes very difficult in a fast moving and dynamic economic context that has been called ‘hyper-competitive’ (D'Aveni, 1994). While isolating the effects of an intervention program in this context can be exceedingly difficult, this does not mean there is no effect.
Timing also plays a role in the cost effectiveness of interventions. It is well known in design science that the cost of a change increases dramatically throughout a development project becoming maximal during the actual implementation phase (Miles and Swift, 1998). This has been suggested to be the case for ergonomics interventions as well with implementation in running systems costing 5-10 times more than in the early design phase (Alexander, 1998). Thus the cost effectiveness of interventions aimed at retrofitting existing systems will tend to be compromised right from the start.

Interventions in the design phase however pose even more problems for evaluation – how can we ‘prove’ a benefit from decisions made that have no counter-example in reality? Comparison relies on available models and ‘virtual ergonomics’ tools an important area for further research and development.

Research Methods - Intervention studies rarely have good quality in both the evaluation and the intervention itself (Volinn, 1999). It seems as if the more rigorous the scientific evaluation is, the weaker the effect that is found (Griffiths, 1999; Karsh et al., 2001). Studies using experimental approaches with randomisation and control groups can increase the believability of results. These studies however, rarely cover the kind of broad multifactor interventions that are suggested to provide the best opportunity for successful outcomes (Westgaard and Winkel, 1997; Karsh et al., 2001; Hignett, 2003; Hignett et al., 2003; Silverstein and Clark, 2004; Hartman et al., 2005). Furthermore, the reductionist strategies of attempting to understand systems by detailed studies of individual components, are inappropriate for complex systems with multiple levels of interaction, relations that change over time, and many uncontrolled variables (Cronbach, 1975; Griffiths, 1999; Karsh et al., 2001) – as is the case with organisational level WEI efforts. The separate system components studied cannot be combined linearly to recreate the whole (Checkland, 1985; Skyttner, 2001). Thus studies of multifactor interventions in complex systems creates a challenge for evaluation – a challenge that may in part be addressed by the use of a series of intermediate outcomes each shedding some light on the overall situations (Cole et al., 2003). The training effects research provided an example of a higher level approach trying to evaluate effects at the industry
level – an approach that might be usefully adapted to more traditional ergonomics practice.

To focus on evaluation quality and not the intervention quality may be seen as putting the cart before the horse. It risks compromising the uptake of potentially useful ergonomic knowledge into workplaces by creating a negative bias in the literature (Dempsey, 2007). Interventions should engage company stakeholders as leadership and participation are frequently seen as crucial for the success of such programs (e.g. Cohen et al., 1997). Such engagement implies lack of control for the researcher and compromises the appropriateness of experimental approaches. Action research provides methodological approaches for researchers operating collaboratively with firms to improve organisational process (Toulmin and Gustavsen, 1996; Reason and Bradbury, 2001; Ottosson, 2003). There remains a need for research on how interventions at the organisational level can affect health of individuals in the company (Fishman, 1999). While new evaluation approaches are needed it appears unlikely that single studies will ‘prove’ causal effects beyond a reasonable doubt. This implies that other aspects suggesting causality, such as time sequence and consistency will become more important. Consistency, however, should be judged within the context of organisational change programs generally which are said to fail more often than succeed (Clegg et al., 2002; Smith, 2003). Longitudinal designs that can use reference groups as benchmarks could be helpful. Furthermore, since interventions can be expected to yield different results at different times in different contexts, reports should emphasise rich detail of the context, content, and process of the interventions so that crucial judgements can be made as to the similarity, and hence possible transferability, of past cases to current situations and contexts. This represents a move away from the notion of a ‘general’ solution to a ‘smorgasbord’ of possible solutions that have been seen to work (or not) in different contexts and that could be adapted to a new situation according to the needs of the local stakeholders.

Quality Criteria Selection Bias - Our literature review suggests that positive results can be achieved with appropriate intervention and many case studies exist demonstrating this. If, on the other hand, you have extensive methodological criteria such as for randomised
controls then there will be considerably fewer studies with positive results. If reviewers base inclusion criteria exclusively on study (evaluation) quality, including only experimental studies, then well conducted multifactor interventions at the organisational level (intervention seen to have highest success potential) will be excluded, and experimental studies with limited interventions will be included. Such reviews cannot be expected to demonstrate much effect of WEI. That reasoning is to some extent supported by evidence that participative interventions against MSDs have been shown effective when study qualities were rated using 11 methodological strength criteria among which comparison groups and randomization were only two criteria (Cole et al., 2005).

Similarly the results of meta-analyses of available literature vary depending on whether or not ‘grey’ literature is included in the analysis (McAuley et al., 2000). We described the impact of inclusion criteria on results as ‘quality criteria selection bias’ (Eklund et al., 2006).

Similar to the quality criteria selection bias is the so-called publication bias – the tendency for only successful projects to be a) written up and b) published (Dickersin and Min, 1993; Torgerson, 2006). This will tend to overestimate the success of intervention approaches and may be preventing us from learning why some interventions are not successful. Publication bias leaves researchers and practitioners unable to learn from our mistakes and thereby hindered in developing more effective intervention tactics and strategies.

*Findings are not stable* - The tendency we observed in the literature for older biomechanical intervention studies to have more substantial and positive results than newer intervention studies might be related to any one of a number of hypotheses. The obvious and severe risks in older workplaces might have been improved leaving only issues that are more difficult to intervene upon. Risk may have shifted from high-load biomechanical mechanisms to more psychosocial and low-level sustained loading mechanisms as workplaces developed (Wells et al., 2007). Work rotation has become more common and this spreads risk around the workforce – diluting the effect of a change and reducing variance in the study population. Notions of what constitutes ‘science’ in reporting may also have changed in the last decades – thus changing our
perception of effectiveness. Returning again to systems theory; workplaces are complex systems with dynamic and unstable relationships between system elements that pose moving targets for researchers. Solutions that work today, therefore, may not be appropriate tomorrow and the design of good working environments may demand ongoing development as the organisation itself adapts to changing social, technical, and economic contexts. This may explain why interventions may work in some contexts but not others.

4.0 Conclusions

Our review found that, from a reductionist perspective demanding traditional experimental methods, there is no certain proof that biomechanical interventions have an effect on MSDs. It seems particularly difficult to reduce work related MSD symptoms once these have begun. There are however many case studies in both ‘white’ and ‘grey’ literature that report positive effects of these interventions. There seems to be general agreement that multifactor interventions have better chances at success while single factor interventions can rarely demonstrate a sizeable impact. The results of reviews eliminating studies of complex interventions, that are not amenable to randomised trials, are subject to quality criteria selection bias.

There is some evidence that work environment investments are profitable at the organisational level – even if methodological problems remain. Economic effects appear to come primarily from quality and productivity improvements with lesser contributions from reduced sickness absence costs.

Conclusions from published literature reviews appear to change over time. Reviews of biomechanical interventions suggest that the effects have moderated over time. It is rare for studies to have both high quality interventions and high quality evaluations. Multifactor interventions are much more difficult to evaluate but are widely believed to pose the best opportunity for successful improvement in the working environment.
Interventions in the working environment appear to have positive measurable effects in some cases but not others. There is a need for new investigation strategies that are suitable for complex systems and for practical and well conducted multifactor interventions, since these are not usually suitable for evaluation by traditional experimental approaches. There is also a need to move beyond studies of the individual employees towards interventions aimed at the organisation level. Action research and large sector level studies provide examples of research approaches that may help address these problems.

Acknowledgements
This work has been supported by the SMARTA Theme of the Swedish National Institute for Working Life, the Research Council of the Ontario Workers’ Safety and Insurance Board (WSIB), and the Canadian Natural Sciences and Engineering Research Council (NSERC).

REFERENCES


of the neck, upper extremity, and low back., US Department of Health and Human Services, National Institute for Occupational Safety and Health, Cincinnati


Ramazzini, B. (1700) *De morbis artificum diatriba*, Typis Antonii Capponi, Mutinae


