European Academy
Of Occupational Health Psychology
Conference Proceedings Series

Bundesanstalt für
Arbeitsschutz und
Arbeitsmedizin

Occupational Health Psychology:
Flexibility, Quality of Working Life and Health

Proceedings of the Fifth European Conference

of the European Academy

of Occupational Health Psychology

BERLIN

20-21 NOVEMBER 2003

Edited by:

Sabir Giga
Manchester School of Management, UMIST

Paul Flaxman
Goldsmiths College, University of London

Jonathan Houdmont
Institute of Work, Health & Organisations, University of Nottingham

Michael Ertel
Federal Institute for Occupational Safety & Health, Berlin

European Academy of Occupational Health Psychology
Conference Proceedings Series

Published by I-WHO Publications

The Institute of Work, Health & Organisations
University of Nottingham
Science and Technology Park
William Lee Buildings
University Boulevard
Nottingham NG7 2RQ UK
Tel: +44 115 84 666 46
Fax: +44 115 84 666 25
E-mail: info@ea-ohp.org
http://www.EA-OHP.org

Printed and sponsored by

Bundesanstalt für Arbeitsschutz und
Arbeitsmedizin
(BAuA)
Nöldnerstr. 40/42, D-10317
Berlin
Germany
www.baua.de

ISBN : 0-9539936-3-9
ISSN 1473-0200

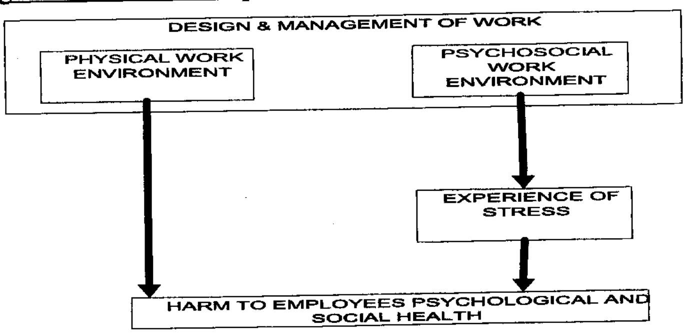
Psychosocial hazards and risk analysis : estimating exposure rates to psychosocial hazards with Latent Class Analysis.

G. NOTELAERS[1] , J.VERMUNT[2] , M. VAN VELDHOVEN[2] and H. DE WITTE[1]
[1] Catholic University Leuven
[2] University Tilburg

From a theory of stress to a framework of risk asessement: calculating psychosocial risks

Contemporary models of occupational health usually describe at least two pathways linking exposure to the hazards of work to their effect on health (Cox, 1993, Cox et al 2000). One pathway traditionally defines occupational hazards in terms of the largely physical-chemical effects of the more tangible hazards of work. The other pathway, which is complementary rather than alternative, is mediated by psycho-physiological and social-behavioral processes, some of which appear to be associated with the experience of stress (Cox, al. 2000).

Fig. 1. The core of dual pathway 'hazard – harm' (Cox, et., al. 2000)

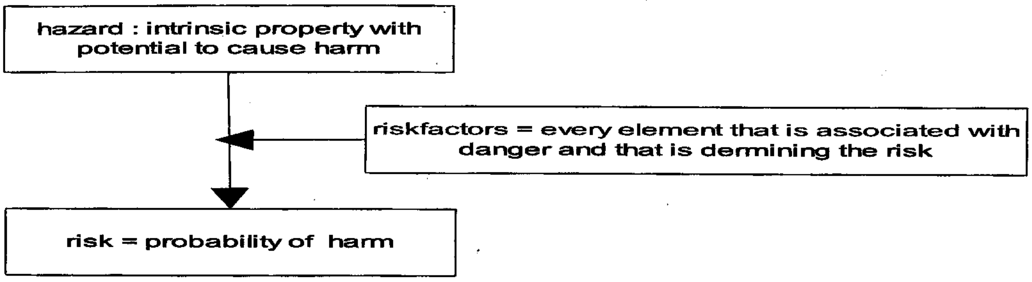


The consequence of this is that the employees' psychological and social health is harmed. In this last step, much of the evidence surrounding harm relates to psychological health and to the risk of cardiovascular disease (Cox et., al. 1996). In this transactional theory of stress, it is important to note that a considerable part of the experience of the stress mechanism is cognitive in nature (Cox, S. et., al. 2000). Central to this, on the one hand, the appraisal theory of Lazarus (1999) is pertinent. On the other hand, coping with these hazards is essential to understand this cognitive process.

A consequence of this indirect pathway is that estimation or calculation of the hazard – risk – harm relationship becomes very difficult. Interesting though, is the direct relationship between physical work environment and psychological and social health. Exposure to physical work environments is often omitted in current stress research. (Levi,1984). The psychological effects of physical hazards reflect not only their direct action on the brain and their unpleasantness but also the worker's awareness, suspicion or fear that they are being exposed to harm. It is the latter which can give rise to the experience of stress (Cox, et., al. 2000)

Lamotte and Van Emelen (1995) integrate the transactional stress model of Cox (1993) in a framework of prevention of hazards. They formulate a conceptual framework for risk assessment that not only deals with the psychological and social health of employees but also with the physical consequences (i.e. injuries).

Figure 2. Relationship between hazard and risk (Lamotte & Van Emelen, 1995)



Conceptualizing risk factors as those factors that can influence hazards, and consequently as factors that contribute to the occurrence of risk, not only withholds the experience of stress as a mediating factor between hazard and harm but incorporates the broader environment in which the job is designed.

For Lamotte and Van Emelen, influencing the risk factors is actually changing the level of exposure to hazards and thus exposure to the risk. In this way, the authors deal with the formal definition of risk (i.e risk = f(hazard,exposure)) and risk assessment : minimizing the risk. In a strict sense this advocates the need for calculating and estimating exposure, and the link with the probable health consequences. However, in order to minimize the risk, identifying the processes and determining the factors that transform hazard into harm are the first steps to be taken. Cox (1993) proposes the use of the control cycle to minimize risks. The control cycle begins with problem recognition and hazard identification. These must be based on a thorough analysis of the work situation. This must include consideration of the tasks and people involved, of procedures and work organization, and of the work environment and relevant technology. The second step 'assessment of associated risk' should both offer an explanation of and quantify the hazard – harm relationship. After that, the implementation of appropriate control measures is to be conducted.

The fourth step initiates the monitoring system. In this step the effectiveness of control strategies should be investigated and questioned. After this step, risks can be reassessed. This is the last step of a recursive system that is designed to ensure continuous improvement of occupational health and safety at work. This control cycle, as a systematic and comprehensive approach to assessing the risks within the work environment, satisfies current legal requirements (Cox, et., al, 2000).

Exposure

The central element in the control cycle is the assessment of hazards. For Cox (1993) this step should be able to explain and to quantify the hazard-harm relationship. Van Emelen & Lamotte (1995) also stress the need to calculate and / or to estimate the link with the probable consequences.

Cox tries to deal with the quantification in two steps. In the first step, the risk assessment team discusses the degree of exposure to, or the likelihood of, hazards. They can categorize this degree as 'absent', 'low or present', 'medium or very obvious' to 'severe'. In the second step, the degree of harm is assigned next to organizational indices for the estimation of severity of harm.

The result is a 3 by 3 'risk matrix' where two values are assigned to each stressor: first, an exposure estimate, and second, an outcome estimate. When the estimation of the severity of outcome is made, account should be taken of existing control and support measures and their effectiveness. When visualized, Cox (et., al, 1996) proposes a priority of risks to be examined : the highest priority e.q. high exposure; highest severity of outcome, the second highest priority, e.q. high exposure; medium outcome and high outcome; medium exposure and third highest priority e.q. low exposure; high outcome. This way of prioritizing is quite arbitrary and consequently Cox (et., al, 1996) warns it should always be remembered that all such devices and schemes for easily estimating risk are weak. They offer no accurate assessment in absolute terms, and they may be unreliable in their detail. They are, however, a useful focus for thinking systematically about the risks and also for subsequent decision-making. In fact, they are a starting point for risk assessment.

Although Cox and others (Cox, 1993, Cox, et. Al, 2000, Briner, et., al, 2001) warn us to seek good indices that make a strong link between hazards and harm possible, they do not question their methods for estimating the probability of exposure. Singleton's (1987) remark is quite justified: much work in psychology seems to have assumed that the evaluation of risk is a purely cognitive process, largely

concerned with estimating probabilities. But this kind of analysis tends to ignore the context in which decision making takes place, and invariably ignores the wider political issues involved. We should be careful with the judgments that members of the team give. Psychological research has shown that people are particularly bad at estimating probabilities, being influenced by memorable and recent events.
Given the data that have already been collected by means of standardized questionnaires like the GHQ or the VBBA, which try to identify hazards and outcomes, we suggest not only using the risk assessment team to estimate likelihood to exposure, but also to use the data for the estimation of the degree of exposure to hazards. Modern statistical classification techniques can be used to propose data-driven ordered exposure rates of hazard that cannot be suspected of being arbitrary. And if risks are being measured it is equally possible to present a data-driven analysis about probabilities of possible consequences, which were measured. Such an analysis can feed the discussion among risk assessors when 'hard' indicators are absent.

Categorizing exposure with questionnaires

In applied research, the aim is to help company assessments of the weak and strong points regarding psychosocial hazards and the psychological health of employees. As in many research projects, this is done by means of analysis of variance. Negative deviations from the mean are categorized as 'weak' and positive deviation from the mean is considered as 'strong'. The result is a brief picture of how well or bad a company is doing regarding psychosocial hazards and psychological health compared to other companies in the same branch. Combined with an analysis at company level (seniority, shift work, type of contract), at departmental level (departments, occupational position) and at individual level (gender, age, education), this method offers the risk assessment team a way of identifying groups at risk.

A well known technique to identify exposure groups departs from a z-transformation. With a threshold like '$z=<-1.65$' the extremes from the mean are being categorized as highly exposed to the psychosocial hazard. The main advantage of such a threshold is that it has well known properties and is thus easy to communicate. This results in a listing of groups (departments, occupational level) having a higher probability of being exposed to psychosocial and physical hazards and psychological and physical health risks (read: to belong to the most extreme 10%).
Another way to identify groups starts from more complicated statistical techniques that deal with categorical data (Hagenaars, 1994). With latent class analysis (Vermunt, 1994) it is possible to model clusters according to the prevalence / exposure of the latent treat. This means that according to the exposure rate to a psychosocial hazard, one can model homogeneous groups that differ distinguishably from one another (Notelaers, Vermunt & van Veldhoven, 2002).
In the next part, the example of recovery, which is measuring whether people feel exhausted at the end of the working day, demonstrates that a statistical measurement model can be used to calculate exposure rates. The accuracy of this model will be confronted with the thresholds derived from a z-transformation of the original Mokken-scale.

Data

These issues will be addressed with the data used in earlier research investigating the cross-cultural equivalence of the measurements of the psychosocial hazards (Notelaers, Vermunt, De Witte, 2003). The data were collected in Belgium between 1999 and 2002 in companies or organizations that used the VBBA (van Veldhoven, 1996) as a diagnostic tool (a self administrated questionnaire) to make an inventory of stressors, organizational facets, strains and facets of well-being. The 9278 observations from the benchmark, have the following composition (in %) : BRANCH : food and textile 12.6, industry 30.7, services 17.6, soc / non profit and government 35.6, health services 3.5; OCCUPATIONAL STATUS : apprentice 0.4, blue collar 20, white collar 49, paramedic / social functions 2.5, lower managerial 12.1, managerial, 11, others 2.1; GENDER : female 46.6, male 53.4; EDUCATION : Primary 3.5, secondary 15.4, high school 27.7, bachelor 26, masters + phd 27.4, SENIORITY : less than 5 years 24.1, between 5 and 10 years 19.6, between 10 and 15 years 16.7, between 15 and 20 years 8.8, 20 years and more 20.7.
However, this is not a representative sample of the Belgian working population. Within the context of our study, i.e. to estimate exposure to psychosocial hazards, this is no problem.

Identifying exposure rates in LCA : the application

This article will try to demonstrate that the use of LCA enables researchers to construct more or less homogeneous groups according to their exposure to a psychosocial hazard or health risk. From this analysis we can easily derive how many people in a sample are being highly exposed to psychosocial hazards and health risks. To reach this goal, we will use the construction of latent classes and the meaning of these classes. Central to assigning the exposure levels are the probabilities answering 'yes' or 'no' to the questions measuring 'recovery need'. These answers constitute an answering pattern that can be allocated to a cluster regarding the likelihood of that pattern.(Notelaers, van Veldhoven & Vermunt, 2002; Notelaers, Vermunt & De Witte, 2003) This means that the answer pattern determines the exposure rate to a psychosocial hazard / health risk.

Recovery need is a Mokken scale measured with 11 items. The respondents could answer with 'yes' or 'no'. The items are listed below next to the profile plot (figure3). The table of fit-statistics obtained from Latent Gold (Vermunt & Magidson, 2000) reveals which cluster model fits the data.

Table 1. Fit-statistic for the latent trait recuperation needs

Model		L^2 (or LL)	BIC	df (or Npar)	proportional reduction of error	p-value
1	1-Cluster	35383.73	-1702.47	4083		5.8e-4886
3	3-Cluster	6839.554	-30010.5	4057	0.81	9.90E-147
4	4-Cluster	5447.222	-31284.7	4044	0.85	1.90E-45
5	5-Cluster	4957.037	-31656.8	4031	0.86	3.20E-22
10	5-Cluster	4168.67	-32399.8	4026	0.88	0.057
11	4-Cluster	4170.181	-32489.1	4036	0.88	0.069
12	5-Cluster	4129.343	-32430	4025	0.88	0.12

Relying on a rule of thumb used in factor analysis, which states that one should 'stop' extracting a new factor when a considerate drop in eigenvalue with the extraction of the next factor no longer occurs (see true dimensionality derived from the scree plot: elbow), leads to the acceptance of the four cluster model(n°4). The difference in L^2 between the five (n°5) and the four cluster model (n°4) is rather small compared to the difference between the four (n°4) and the three (n°3) cluster model. However, at that point, the lowest BIC is assigned with the five cluster model(n°5). Refining the 4 cluster model by adding more direct relationships between indicators and / or adding more direct relationships between indicators and stratifier brings a model (n°11) that is easier to interpret and has a lower BIC than the five-cluster models (n° 5,10,12) presented in the above table. Entropy R^2, similar to the traditional r^2 measure, is 83%.

By looking at the profile table (not shown) it is possible to overlook the probabilities of the answers to items. However, with dichotomous items a profile plot (fig. 3) is more attractive. The meaning of the cluster can be derived from this profile plot. On the X axis the items are given. On the Y axis the conditional probability is given. This is the probability of answering 'yes' (or 'no') to an item given the membership of a certain cluster. From the plot it is obvious that two clusters are clearly separated. Cluster 1 where the conditional probabilities of agreeing with an item are the lowest and cluster 2 where the conditional probabilities are the highest compared to those of other clusters. Cluster 1, covering almost 40% of the sample, is called *'no exposure to recovery need'*. Cluster 2 covering 23% of the sample is called *'high exposure to recovery need'*. That the two other clusters are more difficult to label can be seen in the profile plot: except for the conditional probabilities of two items the two clusters follow the same pattern: the conditional probabilities of cluster 3 are higher or more or less equal to those of cluster 2. But the items 'by the end of the working day, I feel worn out' and 'Because of my job, at the end of the working day I feel absolutely exhausted' show an extreme opposite pattern. It seems to us that cluster 3 is the translation of a low need for recuperation. The cluster withholds 21% of the sample size. The recovering need of the respondents in cluster 4 is higher, further on towards tiredness. As already pointed out, they have a much higher conditional probability of answering 'yes' to items 2 and 3. Above that, the conditional probability of agreeing with "often, after a day's work I feel so tired that I cannot get involved in other activities" is a little bit higher. That of agreeing with "During the last part of the working day, a feeling of tiredness prevents me from doing my work as well as I normally would do" is the same in cluster 3 '*low exposure to recovery need*'. Cluster 4 can be labeled as '*(medium) exposure to recovery need*'. This cluster accounts for 17% of the observations.

Confronting LCA and other ways to define exposure rates.

Comparing the use of z-values to identify individuals who are highly exposed to recovery need is quite common. Often a z-value of -1.65 serves as a cut-off to identify these individuals. Looking at the table below (table 2), this leads to the conclusion that 4,45 % of the sample have high exposure rates regarding recovery need. However, when the z-values are compared to the latent clusters developed in the previous section, one can conclude that this measure is conservative. Accepting the measurement model constructed with latent class analysis, 24,5% of the respondents are highly exposed to recovery need. Confronting these two ways of classifying individuals reveals that 20% of those highly exposed to recovery need are not detected by the z-value. Targeting the 4.45% with policy measures could neglect 80% of the individuals who also suffer from a high exposure rate.

Table 14. Classification z-treshold / latent cluster analyis

Traditional statistical framework			4 cluster of recovery need (in population %)			
scale value	z-value	%	no exposure	low exposure	moderate exposure	high exposure
0	1.25	17.9	17.9			
9.09	0.97	11.9	11.9			
18.2	0.68	8.37	7.16	0.92	*0.3*	
27.3	0.4	8.34	0.89	5.1	*2.35*	
36.4	0.11	7.17		4.42	2.75	
45.5	-0.17	7.7		*4.59*	3.11	
54.5	-0.46	7.58		*3.75*	**3.83**	
63.6	-0.74	7.04		*2.76*	**3.17**	***1.11***
72.7	-1.02	6.42		*0.32*	**0.27**	***5.83***
81.8	-1.31	7.11				***7.11***
90.9	-1.59	5.99				***5.99***
100	-1.88	4.45				*4.45*

It is interesting to see that the mean of recovery need (40) can be attributed to two clusters, i.e. the low exposure cluster and the moderate exposure cluster. Moreover, the mean does not lie in the no exposure cluster. Does this signify that the mean is 'unhealthy'? Within the moderate exposure cluster people have a probability over .8 of agreeing with the items 'By the end of the working day, I feel worn out' and 'Because of my job, at the end of the working day I feel absolutely exhausted'. This is more than two times higher than the low exposure cluster. The probability of agreeing with 'generally, I need more than an hour before I feel completely recuperated after work' and 'When I get home from work, I need to be left in peace' but disagreeing with 'After the evening meal, I generally feel in good shape' is over .6 for individuals situated in one of the two clusters. Given these probabilities we ought to consider that the mean itself could be 'unhealthy' for more than 40% of the individuals located around the mean (see grey numbers in table 2).

The mean corresponds to people who have a considerate probability of agreeing with the fact that they need time for themselves when they get home in order to recover from the working day. Given this interpretation 10% of the sample size (bold underlined numbers) should be added to the number that is neglected by a cut-off criterion of z-values beneath - 1.65. Thus, in total over 30 % of the sample size is neglected. Relaxing the cut-off to a z-value =< -1 reduces that number to 10%.

But following our line of reasoning and given the fact that a z-transformation sets the mean equal to zero, one could conclude that the 6% (sum of underlined numbers) of respondents of the moderate exposure cluster that have a mean score beneath 40 are falsely assumed to be 'healthy' since interventions start as negative deviations from 0.

Conclusion

Risk analysis as a part of risk assessment for psychosocial hazards is a very tricky domain. What is essential is to conceptualize exposure to hazards in order to calculate the risk. Cox suggests leaving the estimation of exposure and severity of outcome to the risk assessment team. Others use cut-offs to differentiate between not exposed and exposed. These are arbitrary and inaccurate. This article tries to show that the collected data can be used to estimate exposure to the (psychosocial) hazards and health risks. From the answering pattern clusters are constructed which reveal the average probability of being exposed to psychosocial hazards. The results show that there are clearly more than two groups i.e. exposed / non exposed. In this sample the cluster model shows that 23% of the sample is highly exposed to recovery need, 17% is moderately exposed to recovery need, 21% has a low exposure rate of need to recover and 40% is not exposed at all.

The ordered latent class solution reveals even more. It shows that the mean itself, the central tendency measure in classical statistical analysis, can be problematic. The mean recovery need in this Belgian sample is related to an elevated exposure to recovery need and thus is not equal to absence of recovery need. This has severe implications. Traditional analysis of variance to find target groups, or z-thresholds to identify extremes, also turns out to be problematic. When compared to the use of a z-value, the ordered latent classes also show that the former are very conservative and inefficient estimates for identifying the troublesome areas. We are confident that the use of the full range of the distribution with ordered latent class analysis is worth further investigation. Such research should also measure health consequences to find out whether there is a truthful and useful link between exposure and harm with the use of latent class analysis.

References

Cox, T. (1993) *Stress Research and Stress Management : putting theory to work.* Centre for Organisational Health and Development, Department of Psychology, University of Nottingham. HSE Contract Research Report, nr. 61.

Cox, T. Griffiths, A. Cox, S. (1996) Work-related stress in nursing : controlling the risk to health. International Labour Office, Geneva. 1996.

Cox, T. Griffiths, A. Rial-Gonzalez, E. (2000) *Research on Work-related Stress.* Institute of Work, Health & Organisations. Research Report for the European Agency for Safety and Health at Work, Bilboa.

Singleton t. Hovden, J. (ed) Risk and Decisions. Wiley, 1987, 232 p.

Lamotte, J-M. Van Emelen, J. Risico-analyse : de nieuwe pijler van het preventiebeleid. In : *Arbeidsblad*, 19, 1995.

Hagenaars. J. (1994) *Categorical Longitudinal Data. Log-Linear Panel, Trend and Cohort Analysis.* Sage Publications

Magidson, J. Vermunt, J.K. (2001) Latent Class Factor and Cluster Models, BI Plots and related graphical displays. In: *Sociological Methodology*, 31, 223-264.

Notelaers, G. van Veldhoven, M. Vermunt, J. (2002) *Towards a methodology for counteracting stress at work. Risk analysis with a standardized questionnaire.* Paper presented at the XVIth World Congress on Safety and Health at Work. Vienna.

Notelaers, G. Vermunt, J. De Witte, H. (2003) *Stress in a multilingual environment. Implications for the validation of standardized questionnaires.* Paper presented at the Conference "Work, Stress and Health: New Challenges in a Changing Workplace", The American Psychological Association (APA). Toronto.

Van Veldhoven, M. (1996) *Psychosociale arbeidsbelasting en werkstress.* Swets en Zeitlinger. Lisse

Vermunt, J. Magidson, J. (2000) *Latent Gold.* User's Guide. Statistical Innovations, Belmont.

Vermunt, J.K. Magidson, J. (2002) Latent Class Cluster Analysis. In : Hagenaars, J.&McCutcheon (eds), *Applied Latent Class Analysis*, Cambridge : Cambridge University Press, 89-106.

Vermunt, J.K. (2001) The use restricted latent class models for defining and testing nonparametric and parametric IRT models. *Applied Psychological Measurement*, 25, 283-294.