National Standard vs. Corporate-Internal Ergonomics Evaluation - an Industrial Case Study

Cecilia Berlin [a,*], Roland Örtengren [a], Dan Lämkull [a, b], Lars Hanson [c]

[a] Department of Product and Production Development, Chalmers University of Technology, SE-412 96 Göteborg, Sweden

[b] Volvo Car Corporation, Manufacturing Engineering Dept. 81121, SE-405 31 Göteborg, Sweden

[c] Department of Design Sciences, Lund University, SE-221 00 Lund, Sweden

Abstract

In Sweden, industrial corporations are required to actively monitor and improve the working environment for their employees according to the provision AFS 1998:1, a document that offers guidelines for maintaining a healthy physical working environment, chiefly by ensuring healthy working postures. This provision was purposely formulated in a very general, non-specific manner in order to be relevant for a large variety of industries.

At certain industrial corporations, corporate-internal procedures for evaluating and monitoring the ergonomic status of the workplace have been developed, usually in order to take consideration of the specific conditions of that industrial application. The protocol, method, execution and presentation of the evaluation results may vary. Also, the decision regarding which personnel should carry out the evaluation becomes a poignant question.

In this article, which describes a Swedish industrial case study in the automotive sector, a corporate-internal method with a highly specified input protocol was used to evaluate the ergonomic status of a production line. At the time, the evaluation was carried out by factory personnel with specific training in using the method. Months later, the same factory segment was re-evaluated by two professional ergonomists from an occupational health service, this time using the AFS 1998:1 provision as a basis.

The article makes an attempt to compare the two methods, finding some similarities and also some interesting differences. However, the main research questions that arise from this work concern the difficulties of choosing the right evaluation method for a large, complex industrial system; not only do different evaluation methods tell us different things at different levels of detail, but they also require very different competencies from the persons who perform the evaluation.

Keywords:

Comparison study, Physical ergonomics, Ergonomics evaluation, National standard

* Corresponding author. Tel: +46-31-772-1290

E-mail address: cecilia.berlin@chalmers.se (C. Berlin)

1. Introduction and Aim

Scientific articles that address corporate-internal or national standard ergonomics evaluation procedures are few and far between. Studies comparing scientifically developed and validated ergonomics evaluation methods in industrial settings have been performed (Drinkaus et al., 2003; Bao et al., 2006; Jones and Kumar, 2007), but the reality is that industrial corporations sometimes develop and use their own procedures for evaluation or, as in this case, use a national provision as pass-or-fail criteria.

This article describes a retrospective corporate case study at Volvo Car Corporation (VCC). Comparing a corporate-internal and a national standard-based procedure with each other illustrates the consequences of changing between vastly different evaluation models – such a change affects output data, which in turn has consequences for personnel who require a specific type of data to fulfil the objectives of their work with ergonomics. This article explores whether there are differences in how the two evaluation procedures identify and rank ergonomic problems, and whether there are any evident patterns in the evaluation results for either procedure.

2. The Studied Evaluation Procedures

Both procedures result in a 'three-zone-rating' for each work station, labelling work stations as *green / yellow / red*.

BME is a Volvo-internal, quantitatively based observation method developed by Amprazis (2005) that identifies risks for musculoskeletal disorders in production (Törnström et al., 2008). Its input structure is a highly specified protocol with rating scales based on specified boundary conditions. BME was developed for use by production personnel in teams consisting of one Local Manufacturing Engineer and one Worker Safety Representative working at the assembly line being analysed, thereby having first-hand experience of performing the work tasks (Björk, 2006).

The Swedish Work Environment Authority issues legal documents regulating corporate work environments. In 1998, the SWEA released the provision AFS 1998:1 (abbreviated AFS-98 here), targeted specifically at physical ergonomics. It is stated explicitly in the provision that an employer is responsible for continually maintaining a healthy workplace for the employees. However, the guidelines are intentionally simplified to increase the scope of application. For this reason, extensive ergonomics knowledge and more detailed evaluation models are needed (and recommended) for more in-depth investigation (AFS, 1998).

3. Method

First, the two evaluations were compared quantitatively on their lowest-common-denominator output data: red-yellow-green ratings. Statistical treatment of the data was comprised of crosstabulation, followed by an assessment of agreement inspired by the quantitative approach described by Bland and Altman (1986), who suggest a nonparametric approach for measuring agreement between two measurement methods. The method is used for comparing agreement between two quantitative measurement instruments or methods, by examining how much the measurements differ from an ideal 'mean' value. This makes the approach suitable for comparing measurements lacking a verifiable 'true' value (e.g. a rating)[1]. Pairwise

[1] An example would be using the Bland-Altman approach to determine whether one instrument for e.g. measuring blood cell counts can be replaced by another. An alternative application could be comparison of two different observers evaluating the same observation to check for inter-rater agreement.

comparisons of cases reporting the Shoulder, Hand, Back and Neck categories were performed, as these were the most numerous body segments in occurrence.
In order to gain insight regarding the historical background, the intended users and practical issues concerning use of either procedure, two semi-structured group interviews were also carried out with a) a group of Pre-production ergonomists who were in charge of factory ergonomics evaluations before the introduction of BME, and b) the OHS ergonomists who are currently in charge of evaluating the factory workstations using AFS-98.

4. Results

4.1. Quantitative comparison

Below is a comparison table of how the two procedures rated stations as being red, yellow or green.

Table 4.1.1: Comparison of green – yellow – red assessments for BME and AFS-98. Numbers in boldface indicate agreement.

AFS-98 ratings vs. BME ratings – Green/Yellow/Red comparison

		BME			Total
		green	yellow	red	
AFS	green	**47**	18	2	67
	yellow	33	**50**	2	85
	red	3	8	**0**	11
Total		83	76	4	163

The agreement between the two procedures was studied on an individual station-level using Bland and Altman's (1986) approach. In the following plot (Fig. 4.1.2), where red =3, yellow = 2 and green =1, each 'step' of difference is counted as 1. The chart is generated by calculating the difference as the BME station rating minus the corresponding rating from AFS-98. Each mark signifies a work station, in sequence as they appear in the factory [2].

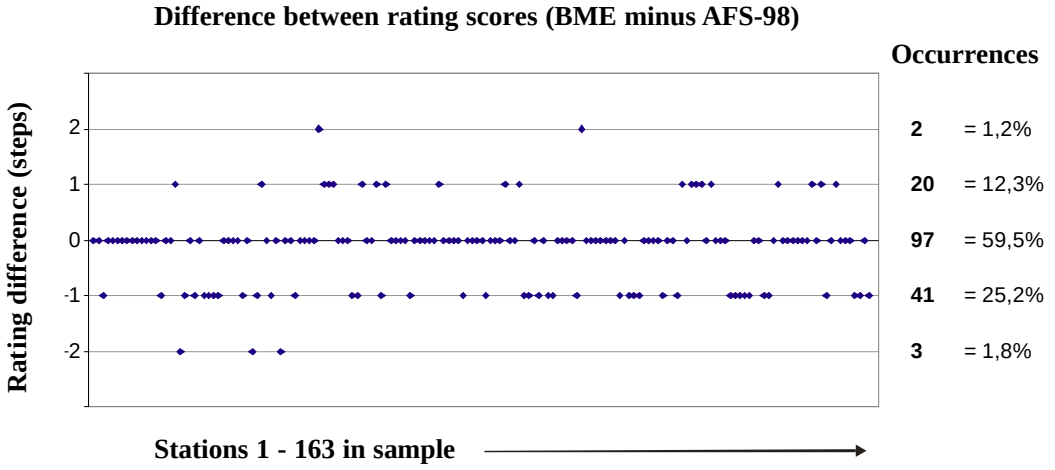

Fig. 4.1.2: Plot of difference between ratings over workstations – each mark signifies one individual station. Negative values indicate a 'more severe' rating from AFS-98 than BME.

[2] Since stations have been filtered out of the sample, the order of sequence is relative and not absolute.

The BME protocols reported a maximum of two body segments while the AFS-98 tables reported all body segments considered at risk, in one case reporting as many as five body segments with no evident ranking. Table 4.2 shows the levels of agreement for Back, Shoulders, Hands and Neck.

Table 4.2: BME-AFS-98 comparison of body segment reports, with statistical probability calculations. Numbers in boldface indicate agreement between both methods.

AFS-98 vs. BME – Back, Shoulder, Hands and Neck

		BME			
		BACK	NO BACK	*Total*	Agreement = 64.1%
AFS-98	BACK	**32**	12	*44*	Pearson χ^2 = 11,398
	NO BACK	31	**45**	*76*	
	Total	*63*	*57*	*120*	Two-sided significance level p= 0.001
		SHOULDER	NO SHOULDER	*Total*	Agreement = 67%
	SHOULDER	**59**	24	*83*	Pearson χ^2 = 8,464
	NO SHOULDER	16	**21**	*37*	
	Total	*75*	*45*	*120*	Two-sided significance level p= 0.004
		HAND	NO HAND	*Total*	Agreement = 58.3%
	HAND	**39**	29	*68*	Pearson χ^2= 3,394
	NO HAND	21	**31**	*52*	
	Total	*60*	*60*	*120*	Two-sided significance level p= 0.065
		NECK	NO NECK	*Total*	Agreement = 78.3%
	NECK	**5**	4	*9*	Fisher's Exact test:
	NO NECK	22	**89**	*101*	p= 0.027
	Total	*27*	*93*	*120*	

4.2. Interviews – Qualitative Comparison

Both interview groups had experience of at least one of the studied evaluation procedures and provided many insights and offered several explanations for some of the tendencies observed in the quantitative data treatment.

The Pre-production Ergonomists have historically performed highly detailed analyses of the assembly line workstations, to a "bolt-and-nut level". Conversely, the main assignment of the OHS ergonomists who performed the AFS-98 evaluation is to rate the work stations on a holistic level. The OHS ergonomists stressed the fact that for practical reasons, they are obliged to keep workstation evaluations to a simplified, holistic level due to time and personnel restrictions.

According to both interviewed groups, the quantitative nature of BME historically led to a tendency to 'hide' red tasks in an assembly sequence in a station that might be rated as green overall. Both interview groups mentioned that there was a role-based conflict of interests inherent within the BME teams; while the Local Manufacturing Engineers' interest was to preferably get an acceptable (green) rating for their assembly line segment, the Worker Safety Representatives had as an objective to expose unacceptable working conditions.

Both interview groups explained the differences in ratings in the quantitative study by emphasizing that the two procedures have different 'algorithms' for arriving at the colour ratings, and also different interpretations of the middle 'yellow' level. Traditionally in the automotive industry, the yellow level has been considered a non-critical level that merely signals room for improvement, but the OHS ergonomists maintained that yellow should be considered a more severe rating according to AFS-98.

5. Discussion

Since this is a situation-specific case study comparison of two not scientifically validated procedures for assessing ergonomics in a workplace, the analysis and approach has been decidedly ad hoc and explorative. Due to the nature of the data collection and the purpose of the procedures, we cannot assume a normal, or indeed any specific type of distribution, of the results or the differences between them. It is more sensible to expect that the specific problems at each assembly line segment should influence the ratings from both procedures in the same way, as they occur sequentially. For this reason, Bland and Altman's (1986) approach for measuring agreement between methods seemed an attractive alternative and was helpful in visualizing the differences.

Tendencies towards a particular colour rating may in each individual occurrence be explained by factors such as method boundary definition differences, underlying work objectives of the personnel performing the evaluations, local changes to the assembly line between evaluations, etc. One point that was brought up over and over in the interviews is that the historical definition of the yellow category at Volvo (prior to the introduction of BME) has been 'acceptable', to the point where yellow has on occasion been considered a target value for some workstations. In contrast, the OHS ergonomists view a yellow rating from AFS-98 as far more severe. This has and may continue to cause misunderstandings between professional groups.

6. Conclusions

- Changes in ergonomics evaluation affect the time consumption for data collection, the level of input detail, the desired or required level of competence (ergonomics-related and industry-specific) of the observers, and the nature and usefulness of the output data for different stakeholders (production developers, OHS personnel, production personnel).
- The main consequence of having changed from BME to AFS-98 has been a change in problem-solving focus (partially due to the shift in personnel allocation for the assignment).
- The lack of agreement between BME and AFS-98 on the red-red level is explained to a great extent by the difference in boundary conditions for red and yellow between the two procedures.
- The shift from a detailed technical focus to a holistic level is a rationalization issue – the OHS ergonomists state that they lack the time to complete their contractual assignments at the same high level of detail as was done previously.
- The long-term results of this change are that the ergonomics focus has shifted, and that different professional groups within the organization have had their supply of data for continued work altered, which in turn alters the discussion between the different professional groups on ergonomics issues in the production system.
- It is recommended that discussions on a managerial level should be initiated to ensure that the interpretation of rating levels is understood by all involved professional groups and that their output data needs are met.

Acknowledgments

Many thanks to Oscar Person, Marita Christmansson, Lars-Ola Bligård, Annki Falck and Dan Högberg for constructive comments during the the progression of this study.

References

AFS. ERGONOMICS FOR THE PREVENTION OF MUSCULOSKELETAL DISORDERS - Provisions of the Swedish National Board of Occupational Safety and Health on Ergonomics for the Prevention of Musculoskeletal Disorders, together with the Board's General Recommendations on the Implementation of the Provisions. Solna, Sweden: Arbetarskyddsstyrelsen; 1998.

Amprazis J. BME. Beräknings Modell Ergonomi Volvo Car Torslanda, utg. 03.1, 2005-01-10 (internal Volvo BME education material, in swedish); 2005.

Bao S, Howard N, Spielholz P and Silverstein B. Quantifying repetitive hand activity for epidemiological research on musculoskeletal injuries—Part II: comparison of different methods of measuring force level and repetitiveness. *Ergonomics* **49** (4) 2006; p. 381–92.

Björk E. The BME Model of Calculation and its Connection to Quality and Productivity. LITH-IKP-EX--06/2349—SE. Linköping: Linköping University (Master Thesis in Swedish; abstract in English); 2006.

Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet, 1 (8476) 1986; 307-10.

Drinkaus P, Sesek R, Bloswick D, Bernard T, Walton B, Joseph B et al. Comparison of ergonomic risk assessment outputs from Rapid Upper Limb Assessment and the Strain Index for tasks in automotive assembly plants, Work 2003 21: 165–72.

Jones T, Kumar S. Comparison of ergonomic risk assessments in a repetitive high-risk sawmill occupation: Saw-filer. International Journal of Industrial Ergonomics 2007; 37, Issues 9-10: p. 744-53.

Törnström L, Amprazis J, Christmansson M, Eklund J. A corporate workplace model for ergonomic assessments and improvements. Applied Ergonomics 2008; 39 (2): 219-28.