EHF Audits: State of the Art and Lessons Learned

Colin G. Drury[1(✉)] and Patrick G. Dempsey[2]

[1] Applied Ergonomics Group Inc., Boulder, CO 80304, USA
[2] National Institute for Occupational Safety & Health, Pittsburgh, PA, USA
pbd8@cdc.gov

Abstract. An ergonomics/human factors (EHF) audit is a methodology for regular review of the fit between people and their working systems. Its objective is to provide proactive guidance on problems or good practices, so that actions can be taken to improve the EHF of a work system, often with an emphasis on safety. An EHF audit can be an appropriate measure of the performance of the EHF function. The objective of this paper is to review the current state of EHF audits from a variety of EHF perspectives and domains, so as to address their value and shortfalls. Audits in the literature and in EHF practice were reviewed (Drury and Dempsey 2020) by considering the audit's objective structure and typical questions, as well as noting the balance between breadth, depth and application time. While much of that review concentrated on the data collection instrument, often a checklist or questionnaire, details of how the audit was to be used were noted. This included the sampling scheme and how results of the audit data collection were to be analyzed and presented to management and workforce. The current paper extends the review to lead to lessons learned that can be applied to future audit systems.

Keywords: Audit · Safety · Lessons learned

1 Why Audit EHF?

EHF audits stand beyond evaluations of individual workplaces or functions to provide input into the state of the level of EHF implementation within a larger system. As such, an EHF audit can provide reliable and valid input into the evaluation of both the organization's level of EHF effectiveness and of the specific functions within the organization charged with EHF design and implementation.

Outside the EHF community, an audit refers to a careful examination of records, such as financial accounts, to "…be certain that acceptable policies and practices have been consistently followed" (Carson and Carlso 1977, p. 2). As Koli (1994) noted, a financial audit comprises four steps:

1. Diagnostic investigation. Describe the business and highlight areas requiring increased care and high risk.
2. Test for transaction. Trace samples of transactions grouped by major area and evaluate.

N. L. Black et al. (Eds.): IEA 2021, LNNS 219, pp. 339–345, 2021.
https://doi.org/10.1007/978-3-030-74602-5_49

3. Test of balances. Analyze content.
4. Formation of opinion. Communicate judgment in an audit report.

All four have relevance to EHF audits: different audits may be needed for different areas of concern, sampling of tasks is more usual than a complete audit of all tasks, each task must have an analysis, and the final results must be communicated to management for action.

2 Examples of Current Audit Systems

We shall follow our own precepts by referencing only a sample of all published EHF audits drawn from the comprehensive review of Drury and Dempsey (2020) which reviewed 30 current and historical EHF audit systems.. The description is meant to give a depiction of available audits, and also to convey the utility of this particular assessment approach that we believe can be used to develop useful and useable assessment tools.

The so-called International Ergonomics Association (IEA) Checklist was presented at the first Congress of the IEA (see Burger and De Jong 1962). The checklist was designed to cover a range of occupational contexts. This was done by using the concept of functional loads (1. physical, 2. perceptual, and 3. mental) as rows of a matrix and system components (A. worker; B. environment; and C. working method, tools, machine) as columns of a matrix to plan the analysis. The ergonomist determines which cells are relevant for the particular analysis being performed. A checklist was provided for each cell requiring analysis, creating an approach that was modular and efficient. This provides the historical context but also foreshadows the form of more current audits.

A more recent modular audit system was developed for assessing ergonomics of aircraft inspection and maintenance activities (Koli et al. 1998). These audits were computerized as the Ergonomics Assessment Program (ERNAP) as described by Meghashyam (1995). While more context-specific than the approach suggested by Burger and De Jong (1962), ERNAP was designed for inspection and maintenance activities that are nonrepetitive in nature. The approach follows specific tasks assigned to technicians by task cards at the start of each shift rather than assessing a workplace. Extensive task analyses of inspection and maintenance tasks were used to develop generalizable function descriptions of inspection and maintenance work (Drury et al. 1990).

The International Labour Office (ILO) in collaboration with the IEA developed "Ergonomic Checkpoints" (ILO 2010) which are easy-to-use checklists for assessing materials handling and storage, hand tools, machine safety, workstation design, lighting, premises, hazardous substances and agents, welfare facilities, and work organization. There are 132 checkpoints, each of which is a checklist item under one of the categories mentioned. Information on how to correct identified deficiencies are given for each checkpoint. The checkpoints are also freely available as a mobile application. The checkpoints are intended to cover a fairly broad range of work settings and were designed to be easy to understand and use.

More recently, Dempsey et al. (2017) developed an extensive set of audits for three types of mining operations (bagging operations, haul truck operations, maintenance and repair operations). The audits are available from the National Institute for Occupational

Safety and Health (NIOSH) in a paper form as well as a more convenient electronic version, ErgoMine (NIOSH 2016), that is available as a free Android application (https://www.cdc.gov/niosh/mining/works/coversheet1906.html). The three types of operation are all quite different. Bagging operations tend to be repetitive requiring materials handling tasks including palletizing and carrying, while maintenance and repair operations additionally involve non-repetitive tasks performed at various locations on a mine site. Haul truck operations involve equipment access during ingress and egress, driving, and tasks associated with refueling and minor maintenance. Like the Ergonomic Checkpoints, each audit item leads to one or more associated practical solutions. Current work is extending ErgoMine to include more specific checklists to address slips and falls as well as a managerial feature to track recommendations and whether they have been addressed.

3 Lessons Learned from EHF Audit Systems

In any organization, "Lessons Learned" can be important inputs into system improvement, although not without their potential difficulties, (Voit and 2006). In this section we address issues in practical audit system design that have arisen in the examples given above and in Drury and Dempsey (2020).

3.1 Use of Legal vs. EHF Good Practice Standards

There is a temptation to use regulatory considerations as the standards for EHF audits, but this will not often be wholly satisfactory. For example, Dempsey et al. (2017) used information from relevant regulations (Title 30 of the United States Code of Federal Regulations) when developing their audits, but they also used information from sources including injury surveillance analyses, task analysis and laboratory studies to develop audit items and associated remedial recommendations. One laboratory study was initiated due to a lack of specific guidance on the selection of grated walkway materials used at mine sites. Pollard et al. (2015) performed a study to evaluate slip potential of commonly observed walkway materials. Specific recommendations were added to the audit about selecting walkway materials that offered higher slip resistance to prevent slips and falls.

There is no reason that EHF audits should not use a mixture of legal requirements and current best practice as the basis for the standards against which to audit. A similar consideration is applicable to the actual audit measurements. If an audit question asks directly whether a measurement exceeds some standard, there can be a temptation for the auditor to mentally round up or down to give a desired outcome. This tendency can be reduced by having the auditor merely record the actual measurement, which is then compared to the standard at the later analysis stage.

3.2 Breadth, Depth and Application Time

As Drury and Dempsey (2020) noted: "Ideally, an audit system would be broad enough to cover any task in any industry, would provide highly detailed analysis and recommendations, and would be applied rapidly. Unfortunately, the three variables of breadth, depth, and application time are likely to trade off in a practical system." Breadth of an audit can mean either many different workplaces in one domain, or many similar workplaces across domains. Depth means the amount of detail included in the audit, not just "is automation involved?" but more probing items such as "can controls and/or displays change modes automatically?" Application time is clearly an increasing function of both breadth and depth; hence trade-offs are likely in these three aspects of the audit system.

3.3 Data Collection Instrument

Almost all EHF data collection methods have been used as audits: checklists, questionnaires, interviews, group techniques, and archival data. Most published audit systems use checklists or questionnaires, so that standard design considerations for these are directly applicable to help ensure that users of checklists make minimal errors. Recent examples can be found in using checklists for procedure design, e.g., Drury and Johnson (2013). Any data collection instrument needs to be usable by different levels of user familiarity, from repetitive daily use to occasional use. This means that some level of additional help may be needed for the occasional user, such as more detailed embedded instructions within a checklist. These are much easier to include when the data collection instrument is presented on an electronic system rather than as a paper form (see Drury, Patel and Prabhu 2000 for more detail). ErgoMine (NIOSH 2016) includes a number of graphics and parameter definitions such as an illustration pointing to the third metacarpophalangeal joint to assist users with measuring "knuckle height" for materials handling tasks.

This brings up the issue of developing more automated instruments than the traditional paper-and-clipboard. Electronic systems work well in areas such as procedure design (Pai 2003) and form the basis of a number of recent audits of mining tasks (Dempsey et al. 2017). Electronic systems can easily include branching logic if certain items are only completed based on the response to other items: This is much more convenient to the user and reduces the chances of error. The results can also be more easily stored and retrieved should the auditor wish to compare results across time.

3.4 Sampling Scheme

In any sampling, we must define the unit of sampling, the sampling frame and the sample choice technique. For an EHF audit, the unit of sampling is not as self-evident as it appears, but a good start is to use the natural unit of the job that is composed of a number of tasks. Note however that one person performing one simple set of tasks is no longer the norm in most developed economies. Teams often rotate between tasks or workplaces, and many jobs have rather ill-defined tasks as they move from direct action to system supervision.

Definition of the sampling frame is more obvious as it derives from the scope of the audit, whether a section, a department, a whole plant or perhaps a geographically extended area for on-site maintenance jobs. Organizations often have defined responsibilities by location or region for ergonomics and safety, and the sampling frame can be chosen to be consistent with these 'natural' organizational boundaries.

Choice of the sampling technique is also rather straightforward and well-covered in EHF texts. Typical concepts are Random, Stratified Random, and Cluster sampling, although in practice a "sample of convenience" is a tempting alternative. This latter strictly prevents statistical inference from the sample and should therefore be avoided. In particular, if the choice of which tasks to observe is left to the audit user on the ground, it may be easy to obtain biased (or at least non-random) results if the auditor has some idea of which jobs may give "better" or "worse" EHF outcomes. Again, the solution is randomness in sampling unless *every* task in the sampling frame is to be included.

3.5 Data Analysis and Presentation

A comment is warranted concerning how data are analyzed and interpreted. Following Koli (1994), Dempsey et al. (2017) designed their audits so that users measured rather than assessed aspects of tasks. Certain dimensions and weights are entered, and the application then interprets the data entered using coded logic to provide the interpretation of whether a remedial recommendation was warranted. This simplifies the auditing task and eliminates the potential reduction of reliability and validity that a judgement could introduce. This approach may increase the complexity of analyzing the results, but the logic can be coded once into a computerized format that will ultimately make it easier for future end users.

Whether a simple checklist or a more extensive modular audit system, there are two levels of analysis: immediate action to improve the specific task and analysis across the whole sample to assess EHF performance level. The first analysis of collected data should focus on presenting the user with a set of actionable and feasible solutions to address the ergonomics or safety deficiencies suggested by the audit. In other words, the audit system should be designed with the user in mind, just as with any system or tool. This allows the user to focus on implementing solutions rather than merely examining and collating the data. At the sample or system level the objective is to integrate and summarize the collected data across tasks or workplaces. Care is needed at this level to resist oversimplification, e.g., by managers demanding a single number for EHF effectiveness which may disguise uneven levels of EHF effectiveness across the sampling frame.

Two of the audit systems described earlier – ErgoMine and Ergonomics Checkpoints – provide users with recommendations for each item where a deficiency is noted. ErgoMine (NIOSH 2016) provides the user a list of recommendations specific to their responses. In order to increase feasibility, several alternatives are given when there are a range of solutions (e.g., automatic palletizer as ideal to eliminate palletizing, versus rotating lift tables to reduce biomechanical stresses while palletizing). The Ergonomic Checkpoints are not as context specific; therefore, the approach utilized was to provide several alternatives that have potential to cover a variety of workplaces. For example, the checkpoint on whether mechanical devices are used for lifting, lowering and moving heavy materials (Checkpoint 9 under Materials storage and handling) shows examples of a portable gantry, hydraulic floor crane, an overhead gantry, and manually powered lift device.

3.6 Audit System Reliability and Validity

For an audit methodology to be of value, it must have demonstrated levels of validity, reliability, sensitivity, and usability. We have covered usability above, but the other three parameters are all critical aspects of proving that an audit system can meet its goals. These measures are traditionally covered in methodology texts, e.g., Wilson and Corlett (1990). Validity is perhaps the most difficult to demonstrate for an audit system, although several studies in Sect. 2 have had their validity measured. Content and construct validity were carefully assessed for the ErgoMine (NIOSH 2016) system (see Dempsey et al. 2017). A retrospective predictive validity for an audit of aviation maintenance was established with some effort by Hsiao et al. (2013a, 2013b). Reliability shows how well a measurement device can repeat a measurement on the same sample unit, e.g., multiple auditors assessing the same workplace or task. As is well known, reliability sets an upper limit on validity and so is an important measure, as well as being relatively simple to assess. "Sensitivity defines how well a measurement device differentiates between entities" (Drury and Dempsey 2020). Sensitivity is not just the precision of the measurement (number of significant figures) but the ability to detect small actual changes in EJF conditions.

4 Conclusions and Recommendations

Auditing has a long history in ergonomics going back to the first IEA Congress (Burger and De Jong 1962), and there have been a number of successful implementations in sectors ranging from mining to transportation. While the research required to properly develop and evaluate audit systems can be significant, audits can be easily implemented using a variety of approaches from pen and paper to applications implemented on mobile devices. A recent survey of professional ergonomists (Lowe et al. 2019) indicated that a high percentage of ergonomists in a number of countries use observation-based assessment tools. Given the robustness of auditing, we believe there is potential for auditing to solve a number of needs for tools to assist ergonomists.

Like many ergonomics assessment tools, there are limited data on validation of longer-term use of audits and the effect on outcome measures such as productivity, errors or injury rates. Mobile applications can provide data on factors such as where mobile apps are used, how many times they are used, and how many items were completed, but these data do not provide information on the degree of implementation of findings or the downstream effects of implementing recommendations. Additional studies will be needed to better understand these outcomes.

Disclaimer: The findings and conclusions in this paper are those of the authors and do not necessarily represent the official position of the National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention. Mention of any company or product does not constitute endorsement by NIOSH, CDC.

Patrick Dempsey, co-author of this Contribution, is an employee of the US Government.

References

Burger, G.C.E., de Jong, J.R.: Aspects of ergonomic job analysis. Ergonomics **5**, 185–201 (1962)

Carson, A.B., Carlson, A.E.: Secretarial Accounting, 10th edn. South-Western, Cincinnati (1977)

Dirken, J.M.: An Ergonomics Checklist Analysis of Printing Machines, vol. 2, pp. 903–913. ILO, Geneva (1969)

Dempsey, P.G., Pollard, J., Porter, W.L., Mayton, A., Heberger, J.R., Gallagher, S., Reardon, L., Drury, C.G.: Development of ergonomics audits for bagging, haul truck, and maintenance and repair operations in mining. Ergonomics **60**(12), 1739–1753 (2017)

Drury, C.G., Dempsey, P.G.: Human factors and ergonomics audits. In: Salvendy, G., Karwowski, W. (eds.) Handbook of Handbook of Human Factors and Ergonomics. Wiley, Hoboken (2020)

Drury, C.G., Johnson, W.B.: Writing aviation maintenance procedures that people can/will follow. Proc. Hum. Factors Ergon. Soc. Annual Meet. **57**, 997–1001 (2013)

Hsiao, Y.L., Drury, C., Wu, C., Paquet, V.: Predictive models of safety based on audit findings: part 1: model development and reliability. Appl. Ergon. **44**(2), 261–273 (2013)

Hsiao, Y.L., Drury, C., Wu, C., Paquet, V.: Predictive models of safety based on audit findings: part 2: measurement of model validity. Appl. Ergon. **44**(4), 659–666 (2013)

Koli, S.T.: Ergonomic Audit for Non-repetitive Task," unpublished M.S. thesis, State University of New York at Buffalo, Buffalo (1994)

Koli, S., Chervak, S., Drury, C.G.: Human factors audit programs for nonrepetitive tasks. Human Fact. Ergon. Manuf. **8**(3), 215–231 (1998)

Lowe, B.D, Dempsey, P.G., Jones, E.: Ergonomics assessment methods used by ergonomics professionals. Appl. Ergon. **81** (2019). https://doi.org/10.1016/j.apergo.2019.102882. Article 102882

Meghashyam, G.: Electronic ergonomic audit system for maintenance and inspection. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 39, pp. 75–78. Human Factors and Ergonomics Society, Santa Monica (1995)

NIOSH: ErgoMine 1.0. By Dempsey, P.G., Pollard, J.P., Porter, W.L., Mayton, A.G., Heberger, J., Reardon, L., Fritz, J.E., and Young, M. Pittsburgh, PA: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, Mobile App (2016)

Pai, S.: Effectiveness of Technological Aids in Checklist Use. unpublished Master's thesis, State University of New York at Buffalo, Buffalo, NY (2003)

Pollard, J.P., Heberger, J.R., Dempsey, P.G.: Slip potential for commonly used inclined grated metal walkways. IIE Trans. Occup. Ergon. Human Fact. **3**(2), 115–126 (2015)

Voit, J.R., Drury, C.G.: Supporting Vicarious Learning with Collaborative Lessons Learned Programs. IEEE-SMC Part 1. **36**(6), 1054–1062 (2006)

J. R. Wilson, J.R., Corlett, E.N. (eds.): Evaluation of Human Work. Taylor & Francis, London (1990)