CHAPTER 39

HUMAN FACTORS AND ERGONOMICS AUDITS

Colin G. Drury
University at Buffalo: State University of New York
Buffalo, New York

Patrick G. Dempsey
National Institute for Occupational Safety & Health
Pittsburgh, PA

1 PURPOSE

This chapter has two interrelated aims. First, we examine the human–machine tasks of inspecting, checking, and auditing to provide understandings that can guide work design, equipment design, and job aid development. Second, we apply this knowledge to inspecting, checking, and auditing of human factors/ergonomics aspects of human–machine systems. As part of this aim, we provide a detailed review and worked example of recent human factors/ergonomics audit programs. Throughout the chapter examples are given from a wide range of domains, from product usability audits, through aviation preflight checklists, to inspection of products for quality assurance.

2 INSPECTING, CHECKING, AND AUDITING

The idea of inspecting is as old as civilization. In the Sumerian epic *Gilgamesh*, almost 5000 years old, the narrator invites the reader to examine the quality of the walls of Uruk, built by Gilgamesh, the king (*Gilgamesh*, Tablet 1):

> Look at its wall which gleams like copper (?),
> inspect its inner wall, the likes of which no one can equal!
> Take hold of the threshold stone—it dates from ancient times!
> Go close to the Eanna Temple, the residence of Ishtar,
> such as no later king or man ever equaled!
> Go up on the wall of Uruk and walk around,
> examine its foundation, inspect its brickwork thoroughly.
> Is not (even the core of) the brick structure made of kiln-fired brick,
> and did not the Seven Sages themselves lay out its plans?"

The essence of the examining function is all there already. Bodily senses (look, take hold of) are used to compare the existing item (wall) with some implied or actual standard (e.g., kiln-dried brick). Inspection can have more formal definitions (e.g., in dictionaries and quality control texts), but a simple definition from Drury (2002, p. 27) provides a reasonable modern start: "The [test and] inspection system determines the suitability of a product or process to fulfill its intended function, within given parameters of accuracy, cost and timeliness."

Inspection is thus a decision function, for management, for the general public, and for an individual. Should the process be stopped before it produces more defects? Does this café meet local standards of cleanliness, so that people can eat there safely? Is this aircraft safe for me to fly in? Does this box of strawberries contain any bad fruit? The most basic decision is a go/no-go decision: Does this item or process fulfill its intended function or not? In practice, the amount of inspection, checking, and auditing needed to reach such a simple conclusion may be considerable. For example, a national safety board may need considerable evidence after a major accident to determine that the system (aircraft, train, ferry, spacecraft) is fit to resume normal operations. As noted above, most inspection decisions are simpler than this.

In principle, inspection can be done by the producer or consumer of the goods or services directly, but this is not always satisfactory. Thus, the person machining a part can determine whether or not it meets standards, or the ultimate end user can inspect the part. Much of the quality revolution from the 1970s onward has been concerned with pushing such decisions back along the production system to ensure that decisions are made at the source (e.g., Evans and Lindsay, 1993; Drury, 1997) to prevent errors from propagating through a production system. Indeed, the preferred solution is to inspect the *process* rather than the *product* to ensure that defects are extremely unlikely to be produced. That is the aim of in-process statistical process control (e.g., Devor et al., 1992). Unlike our example of the box of strawberries, most ultimate consumers are not equipped to be able to inspect or check the complex devices and processes they use in daily living. In an agrarian economy, a farmer could examine a spade produced for him by the local blacksmith (e.g., Jones, 1981) with enough skill and visual/haptic information to reach a reasonable conclusion about whether the design and construction quality met his requirements. In our more specialized economy, consumers cannot make valid quality judgments on automobiles, computers, or even the safety of a local chemical plant because they lack both the specialized knowledge required and access to the inner workings of the product or system.

In such cases, customers rely on the judgment of professional inspectors, checkers, and auditors to make informed decisions. This reliance raises questions of honesty, trust, competence, and human–machine system design. From a sociotechnical systems perspective, inspectors must be seen as independent of producers of goods and services or their findings will not be accepted. For example, in civil aviation the U.S. Federal Aviation Administration (FAA) decrees that airline inspectors charged with checking airworthiness are kept organizationally independent of aviation maintenance technicians, who perform repairs, adjustments, and replacements. This independence can lead to role ambiguity and conflict in their job. For example, McKenzie (1958) noted that inspection is always of people: The inspector is judging the work of others by examining the outputs. Early work by McKenzie (1958), Thomas and Seaborne (1961), and Jamieson (1966) established that social pressures are an important part of the inspector's job and inspectors can at times change their behavior and performance in response to such pressures. The author worked in one factory where the customer returned shipments of product when they could not be used right away (e.g., during a strike) by finding a defect in the shipment. Later, when the customer needed the product, the factory repackaged the same shipments, often receiving complements on their improved quality!

This chapter covers both inspection and auditing, so that we need to establish that auditing and inspection do have commonalities or are indeed both instances of the same systems function and human behavior. Both clearly involve making decisions about the fitness of a product/service (inspection) or a system (auditing) for its intended use, with heavy overtones of protecting the public. However, a more detailed proof of congruence must be postponed until we have considered each process in more analytic detail (Section 4).

3 INSPECTION AND HUMAN FACTORS ERGONOMICS

As noted above, the study of inspection and inspectors is at least 50 years old, although most concentrated on inspection of products in a manufacturing setting. There are many other types, as noted by Drury (2002, 2009):

- *Regulatory inspection*—to ensure that regulated industries meet or exceed regulatory norms. Examples are review of restaurants against local service codes, fire safety inspection of buildings, and safety inspection of workplaces.
- *Medical inspection*—to ensure that a patient receives correct diagnosis of medical conditions. An example is inspection of mammograms (Nodine et al., 2002; Chen et al., 2008).
- *Maintenance*—to detect failure arising during the service life of a product. This failure detection function can be seen in inspection of road and rail bridges for structural determination or of civil airliners for stress cracks or corrosion.
- *Security*—to detect items deliberately concealed. These may be firearms or bombs carried onto aircraft, drugs smuggled across borders, or camouflaged targets in aerial photographs. Examples are X-ray inspection of airline baggage (e.g., Gale et al., 2000: Hsiao et al., 2008; Ghylin et al., 2007). They can also be suspicious happenings on a real-time video monitor at a security station. Law enforcement has many examples of searching crime sites for evidence.
- *Design review*—to detect discrepancies or problems with new designs. Examples are the checking of building drawings for building code violations, of chemical plant blueprints for possible safety problems, or of new restaurant designs for health code violations.
- *Functionality testing*—to detect lack of functionality in a completed system. This functional inspection can often include problem diagnosis, as with checks of avionics equipment in aircraft. Often functional inspection is particularly dangerous and costly, as in test flying aircraft or checking out procedures for a chemical process.

All of these inspection applications have much in common. Indeed, Drury (2003) has proposed a unified model of inspection in the security domain and shows how it can apply to all security inspection systems. This model was hardly new; it was based

Table 1 Generic Function, Outcome, and Error Analysis of Test and Inspection

Function	Outcome	Errors
Setup	Inspection system functional, calibrated correctly, and capable	1.1. Incorrect equipment 1.2. Nonworking equipment 1.3. Incorrect calibration 1.4. Incorrect or inadequate system knowledge
Present	Item (or process) presented to inspection system	2.1. Wrong item presented 2.2. Item mispresented 2.3. Item damaged by presentation
Search	Indications of all possible nonconformities detected, located	3.1. Indication missed 3.2. False indication detected 3.3. Indication mislocated 3.4. Indication forgotten before decision
Decision	All indications located by search measured correctly and classified correct outcome decision reached	4.1. Indication measured incorrectly 4.2. Indication classified incorrectly 4.3. Wrong outcome decision 4.4. Indication not processed
Respond	Action specified by outcome taken correctly	5.1. Nonconforming action taken on conforming item 5.2. Conforming action taken on nonconforming item

Source: Drury (2002).

on function analytical models developed earlier by Drury (1978), Sinclair (1984), and Wang and Drury (1989). A recent incarnation can be seen in Jiang et al. (2003). Table 1 gives a generic functional breakdown of inspection, showing the major functions of inspection with the correct outcomes and errors arising from each function.

The reader is referred directly to Drury (2002) for a detailed consideration of inspection, including automated inspection and test, and a design methodology for "design for inspectability." In the current chapter, only an overview is provided to help understand the context for both checklists and audits. Each function is considered in the order given in Table 1. A single example of safety inspection of a workplace will be used to illustrate each function:

1. *Setup.* In this function, the inspection system is prepared for use. Needed tools, equipment, and supplies are procured, procedures are available to aid the inspector, and the inspector has been trained to perform the task correctly. For a workplace safety inspection, there will be some equipment needing calibration (e.g., psychrometer, air-sampling systems), a written procedure in the form of a checklist or computer program (e.g., Wilkins et al., 1997), and safety inspectors who have undertaken training and often certification. Certification is one way in which the competence and independence of inspectors are maintained, leading presumably to a higher degree of public trust in the inspection process. The SETUP function places demands on the regulatory system to provide the needed antecedents of effective inspection (i.e., sufficient resources).
2. *Present.* Here the inspector (suitably equipped) and the entity to be inspected come together so that inspection can take place. The narrator of *Gilgamesh* urges the reader to "... go up to the wall of Uruk and walk around." Often, this function is purely mechanical: The manufacturing inspector has products arrive and depart on a conveyor, the safety inspector accesses all areas of a factory that could contain indications of an unsafe condition, and the FAA inspector goes to the file cabinets where maintenance records are kept to check that maintenance was performed and signed off correctly. The managerial implications for PRESENT are that inspectors must know the places they need to examine, and the organization being inspected must provide open and timely access to such sites. In safety inspection by a regulatory agency, there is often special legal provision for sites to be made available with little or no prior notice to prevent concealment of safety concerns known to management.
3. *Search.* Here we arrive at the first of the two most important and error-prone functions of inspection. Search is typically a sequential serial process during which the entire item to be inspected is brought under scrutiny piece by piece. The most obvious form of search is visual, and this has a long history of study, going back to the earliest days of human factors (Blackwell, 1946; Lamar, 1960). People (i.e., inspectors) search an area by successive visual fixations where the eye remains essentially stationary. What can be detected during a single fixation is a function of target and background characteristics (e.g., Overington, 1973) with considerable research (e.g., Treisman, 1986) on combinations of target and background conditions favoring rapid, parallel ("preattentive") search. Mathematical models (e.g., Morawski

et al., 1980; Wolfe, 1994) of the visual search process emphasize two features:

a. Over what area around the fixation point is the target detectable in a single fixation ("visual lobe") (e.g., Chan and Chiu, 2010)?
b. How are successive fixations sequenced to achieve coverage of the entire area?

Visual lobe models (e.g., Engel, 1971; Eriksen, 1990) determine how much area can be covered in a single fixation and hence the number of fixations required for coverage. The time to search an item completely varies directly with the number of fixations required and hence the reciprocal of lobe area (Morawski et al., 1980). Successive fixations are determined partly by top-down strategy and partly by bottom-up features of the currently fixated area (Wolfe, 1994). Top-down strategy has been modeled (e.g., Hong and Drury, 2002) as sequential, random, or a mixture (e.g., Arani et al., 1984). It is determined in part by the inspector's knowledge and expectations of where targets are likely. In an inspection context, a key derivation from search models is the *stopping time*, when the inspector decides that enough time has been spent on one item and moves to the next item. Stopping time chosen by inspectors accords quite well with the predictions of optimum stopping models (e.g., Chi and Drury, 1995; Baveja et al., 1996). Stopping time is a physical manifestation of how many resources the inspector (or the system giving inspection instructions) is willing to devote to each item inspected.

Visual search is not the only form of search important in inspection; there is also procedural search, where an inspector goes through a list of places that require examination. Procedural search is used extensively in aviation checklists (see later) where, for example, a preflight inspection of a general aviation aircraft follows a written procedure requiring examination of control surfaces, tires, fuel, structural joints, and so on. For each item on the checklist, the inspector is trained on what defects to look for. An example is fuel, where a small sample of fuel is drawn from a low point in the fuel line to check for water contamination, with water drops appearing as spheres in the fuel sample.

In a safety inspection, the inspector will have a list of key items/areas to search for indications of lack of safety. Inspectors will use their senses to examine each item in this procedural search, often requiring visual search within the procedural search. For example, inspectors must check that guarding is present on moving machinery, a largely procedural task. When they check safety records, such as the OSHA log in the United States, a visual search is required to determine whether all fields have been filled in and signed correctly. Note that in both these examples considerable knowledge and skill are required to understand what would be an indication of a safety violation.

As noted in Table 1, the successful outcome of search is something detected that *could* be a defect. This is known in the nondestructive inspection (NDI) community as an *indication*. Subsequent inspection functions are concerned with how to deal with each indication. Note that if search fails, the indication is missed and subsequent functions cannot proceed. At least in visual inspection there is considerable evidence (Drury and Sinclair, 1983; Drury et al., 1997) that the search function is quite error prone, with only about 50% of defects ever being located as indications.

4. *Decision.* This is the function in which the indication is judged against a standard to determine whether it is a true defect. If inspection is about decision, this function represents the essence of inspection. Decision requires human (or machine-aided) judgment against a standard, so a standard must be prespecified. Standards in manufacturing inspection can come from physical properties (hardness, conductivity, surface finish) that can be measured with appropriate gauging. The decision in such cases of what the statistical quality control (SQC) community calls *variables inspection* can be automated quite simply and is thus rarely an appropriate human task. Not all measurements and standards can be implemented so simply. How do we quantify blemishes in the surface finish of automobile paint work (Lloyd et al., 2000) or corrosion areas on an aircraft fuselage (Wenner and Drury, 1996). These examples of *attributes inspection* require a complex, typically human judgment. Often, signal detection theory (SDT) has been used as a model of this part of the inspection process (e.g., Drury and Sinclair, 1983; Chi and Drury, 1995) although at times it has been misapplied to the overall inspection process, hence lumping search errors with wrong decisions to accept a true defect (e.g., Drury and Addison, 1973). SDT suggests a separation of decision difficulty (discriminability) from bias in reporting/not reporting defects (criteria), thus providing a useful link to different remedial actions, depending on whether the discriminability or criterion needs changing. For example, Drury et al. (1997) found that the decision function of inspection was extremely inconsistent between inspectors, implying the need for better training and job aids in aircraft inspection.

For safety inspection, discriminability represents the decision difficulty. This can be very easy when the implied standard is zero (e.g., *any* missing machine guard is a violation) or more difficult (e.g., how much untidiness represents "poor housekeeping"). The decision criterion reflects the willingness to report. From SDT this is a function of the a priori probability of a defect and the relative costs of the two errors:

- *Miss*: not reporting a true defect
- *False alarm*: reporting an indication that was not a true defect

In safety inspection, the pressures on the inspector can be high. A false alarm can be used by the factory being inspected to bring disrepute on a regulatory agency and on the actual inspector. A miss can lead to an accident, injury, or even a Bhopal-like disaster. Similar pressures exist for aircraft inspectors, security operators, and even the intelligence community. The decision to report becomes even more difficult (from SDT) when the defect found is extremely rare. A recent example is the failure to find a crack in the titanium hub of a jet engine that caused the accident in 1997 at Pensacola, Florida. The inspector, despite years of experience, had never encountered a crack in a titanium hub before. More details on mathematical approaches to the decision function may be found in Drury (2002). An example of applying a mathematical model of the search plus decision functions to security inspection can be found in Ghylin et al. (2007).

5. *Response.* When the decision has been made, the action chosen must be taken. In a manufacturing context, this can be as simple as removing a defective item from the production process or as complex as stopping the process to diagnose the "root cause" of the defect being produced. In more general contexts, the action is often written (e.g., a repair order for a crack in an aircraft structure, written warning of unsanitary conditions in a restaurant, or a safety citation to company management for a missing machine guard).

Although response is a relatively mechanical function, it can be subject to errors. Aircraft inspectors who intend to write up all defects at the end of inspection can forget some defects, surgeons can mark the wrong leg for amputation, and the safety citation can be incomplete. This is one function where computer-based automation can help by making response simple and immediate. For example, Drury et al. (2000) developed a computer-based task card system for aircraft inspection that allowed easy generation of repair orders. Even more could have been done with drop-down menus for fault type (crack, corrosion, etc.) or a search function for the appropriate reference in the structural repair manual.

Throughout this short treatment of inspection, we have seen where and why good human factors/ergonomics practices can reduce error potential. There are also overarching considerations of job design and automation that can have a great impact on errors [e.g., the hybrid automation studies of Hou et al. (1993) and Jiang et al. (2003)]. Again, details may be found in Drury (2002).

The audit and checking activities associated with nonmanufacturing applications can now be placed in a suitable context.

4 CHECKING AND CHECKLISTS

When inspection is too complex to be carried out by an inspector unaided by procedural notes, a job aid is required to lead the inspector through the task. The nonmanufacturing examples given already (e.g., aircraft maintenance, safety inspectors) are typical of those requiring and using job aids. The simplest job aid for any procedural task (e.g., preparation for landing an aircraft) is the checklist. All pilots carry checklists for many complex procedures when a sequence of actions must be performed in a standard order. In addition to the landing preparation noted above, there are checklists for preflight inspection, startup/taxi, pre-takeoff, climb, cruise, postlanding, and engine shutdown. These are typically short laminated paper lists in general aviation or computer-based lists for corporate and passenger jets. Glider pilots use laminated checklists but also use mnemonics, such as "STALLS" for prelanding, where they might not have time to consult a written checklist. Safety inspectors in industry typically use a written checklist of several pages.

If the form of a checklist (paper, computer, mnemonic) can vary, so can the content and structure. Most checklists are used as memory aids for well-practiced tasks, so that they are structured as lists of commands, each of which is relatively tense: "switch both magnetos on: open fuel cock: prime engine for 5 seconds." The user is expected to know which are the magneto switches and which way they move for "on." Users are also expected to understand *why* each action is required so that they have some strategy for recovering from malfunctions (e.g., one magneto not working correctly).

In contrast, more detailed procedures are used for inspection and maintenance of aircraft and spacecraft. These procedures will spell out each step in detail, often with part numbers and numerical settings (e.g., tightening torque), and include warnings and cautions as well as a rationale for the overall procedure. With a computer-based system, detailed procedures can also be viewed as checklists: for example, when the procedure is being repeated after a short interval. Drury et al. (2000) used simple hypertext links to move between the checklist steps and the more detailed procedures in their program for inspection workcards. Much human factors design and evaluation has gone into the physical design of such procedures (e.g., Patel et al., 1994; Chervak et al., 1996; Drury et al., 2000), so in the remainder of this section we concentrate on classic checklists, as they are most often encountered in nonmanufacturing inspection and audit.

Checklists have their limitations, though. The cogent arguments put forward by Easterby (1967) provide a good early summary of these limitations in the context of design checklists, and most are still valid today. Checklists are only of use as an aid to designers of systems at the earliest stages of the process. By concentrating on simple questions, often requiring yes or no answers, some checklists may reduce human factors to a simple stimulus–response system rather than encouraging conceptual thinking. Easterby quotes Miller (1967): "I still find that many people who should know better seem to expect magic from analytic and descriptive procedures. They expect that formats can be filled in by dunces and lead to inspired insights.... We should find opportunity to exorcise this nonsense" Easterby, 1967, p. 554).

Easterby finds that checklists can have a helpful structure but often have vague questions, make nonspecified assumptions, and lack quantitative detail. Checklists are seen as appropriate for some parts of ergonomics analysis (as opposed to synthesis) and are even more appropriate to aid operators (not ergonomists) in following procedural steps. Clearly, we should be careful, even 30 years on, to heed these warnings. Many checklists are developed, and many of these published, that contain design elements fully justifying such criticisms.

Most formal studies of checklist use have been in an aviation context, both in maintenance (Pearl and Drury, 1995) and in preflight inspection (Ockerman and Pritchett, 1998, 2000, 2004). They have also found widespread use in the flight operations side of aviation, with detailed analysis by Degani and Wiener (1990). The Degani and Wiener (1990) study laid the basis for much subsequent work on checklists. They analyzed incident reports from the National Transportation Safety Board (NTSB) and ASRS (NASA's Aviation Safety Reporting System), finding that the main checklist errors resulted from overlooking items following interruptions or distractions, particularly when working under time pressure or toward the end of the working day. Their recommended countermeasures were use of a "challenge and response" operating philosophy, grouping several items together and using a logical flow pattern. In particular, they advocated a "geographical" sequence of steps, good formatting/typography, and that "operators should keep checklists as short as possible to minimize interruptions." They also reviewed then-current technologies that could assist checklist use. An earlier study of checklists for circuit board inspection (Goldberg and Gibson, 1986) also found that a logically organized checklist outperformed a randomly organized one.

Patel et al. (1993) found that during the initial inspection of an aircraft on arrival at maintenance the sequence of tasks in the checklist or workcard did not match the sequence of tasks that aircraft maintenance technicians typically followed. In a related study by Pearl and Drury (1995), questionnaires and videotapes showed that mechanics tended to sequence their tasks using spatial cues on the airplane rather than the order specified on their workcard. The study also revealed that aviation maintenance technicians who performed low-level inspections used spatial locations of tasks to sequence them. In addition, many aircraft mechanics rarely used the checklist and viewed it as an only guide for inexperienced mechanics. Experienced inspectors felt that they had acquired sufficient skill to perform the inspection task using their memory and referred to the checklist only occasionally.

A more recent series of investigations (Ockerman and Pritchett, 1998, 2000, 2004) have examined the relationship between the medium (paper vs. wearable computers) on which the procedure was displayed, the presentations of procedure context, overreliance, and inspection performance for a preflight inspection task. The studies found that inspection performance could be influenced by the presence of procedure context information presented with procedures. The 1998 study also observed that one-third of the participants used their memory and not the task guidance system to perform the preflight inspection. They observed that in some sessions the subjects performed the task from memory and consulted the checklist only to see if anything was forgotten, echoing the Pearl and Drury (1995) findings from maintenance.

Computer applications for checklists were also advocated by Degani and Wiener (1990) and tested in an aviation maintenance/inspection environment by Drury et al. (2000). The latter study measured the impact of a hypertext-based computer program on the usability of work documentation in maintenance and inspection. Based on data collected in 1992–1993 at an airline partner, they concluded that computer-based inspection job aids were effective, although much of their effectiveness was attributed to good job aid design rather than computerization per se. Their task was one of detailed inspection of aircraft structures and used a checklist only as a top-level job aid, with more detailed instructions and data available via hyperlinks.

Major findings of all of these studies should be applicable to audit checklists, despite their somewhat different domains within aviation. Checklists are good job performance aids for repetitive tasks. They involve little explanation of detail or rationale for the sequence of operations, being mainly reminders of the correct sequence, often with facilities for marking (the "check" in "checklist") each item as it is performed to minimize the effect of interruptions or distractions (Degani and Wiener, 1990). The findings include (1) a geographical sequence is probably best; (2) good design principles should be followed; (3) technology can improve checklists; and (4) checklists are not always used, with reliance often placed on memory. The latter finding was reinforced by Wenner and Drury (2000), who note that some people did not read or follow very explicit instructions for performing a task.

Recently, two closely related studies of checklists were performed using a simulated repetitive aircraft inspection task with engineering student participants. The first (Larock, 2000; Larock and Drury, 2003) measured the effects of checklist layout and the number of sign-offs when a task was repeated eight times on eight days. The second (Pai, 2003) examined the use of computer-based checklists under the best design conditions found by Larock and Drury (2003) using the same task repeated over six days.

The first study compared functionally ordered and spatially ordered checklists and also whether each of the 108 items had to be signed off individually or whether they were signed off in 37 logically related subsets. As expected, spatial ordering was better than functional ordering for both accuracy and speed. The number of sign-offs and the checklist layout interacted for sequence errors, where the best combination was spatial ordering and signing off in 37 groups. Over the course of eight daily trials, participants became faster at the task but tended to develop a spatial strategy for either checklist. The second study used this combination to test the efficacy of various computer implementations of the original paper checklist. A personal digital assistant (Palm-Pilot) was used with its built-in application of the "to-do list." A more user-friendly program was written specifically for the task studied and was implemented on the PDA and on a laptop computer. The conclusion was that the three computer implementations did not differ from each other, and all gave better speed and accuracy than those for the paper-based checklist.

These studies reinforce conclusions 1–4 noted above and showed that checklist behavior is not merely an artifact of using aviation professionals as subjects. Checklists emerge as a powerful tool, but one that needs careful human factors design to reach its maximum performance. Outside of aviation operations, there is little evidence that checklists are designed with these human factors findings in mind.

To develop checklists for auditing safety, ergonomics, or human factors, the design principles above should be followed. In addition, checklists need to be validated with actual users to ensure that their content, structure, and format do indeed lead to reliable performance.

5 AUDITING WITH SPECIFIC APPLICATION TO HUMAN FACTORS

When we audit an entity, we perform an examination of it. Dictionaries typically emphasize official examinations of (financial) accounts, reflecting the accounting origin of the term. Accounting texts go further: for example, "testing and checking the records of an enterprise to be certain that acceptable policies and practices have been consistently followed" (Carson and Carlson, 1977, p. 2). In the human factors field, the term is broadened to include nonfinancial entities but remains faithful to the concepts of checking, acceptable policies/practices, and consistency.

As with inspecting, auditing is mentioned in antiquity, at least in current translations: "[H]e who does not pay the fine annually shall owe ten times the sum, which the treasurer of the goddess shall exact; and if he fails in doing so, let him be answerable and give an account of the money at his audit" (Plato, *Laws*, Book VI).

Human factors audits can be applied, as can human factors itself, to both products and processes. Both applications have much in common, as any *process* can be considered as a *product* of a design procedure, but in this section we emphasize process audits because product evaluation is covered in detail in Chapter 50. Product usability audits have their own history (e.g., Malde, 1992), which is best accessed through the product design and evaluation literature (e.g., McClelland, 1990).

Auditing, like inspection, proceeds through a series of functional steps. For example, an audit by a certified public accountant would comprise the following steps (adapted from Koli, 1994):

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.
3. *Test of Balances*. Analyze content.
4. *Formation of Opinion*. Communicate judgment in an audit report.

There are obvious direct parallels with the functions of inspection (Table 1), as noted by Drury (2009): *Diagnostic Investigation* comprises the Setup task, *Test for Transaction* comprises the Present and Search tasks, *Test of Balances* is the Decision task while *Formation of Opinion* is the Response task.

5.1 Need for Auditing Human Factors

Human factors or ergonomics programs have become a permanent feature of many companies, with typical examples shown in Alexander and Pulat (1985). As with any other function, human factors/ergonomics needs tools to measure its effectiveness. Earlier, when human factors operated through individual projects, evaluation could take place on a project-by-project basis. Thus, the interventions to improve apparel sewing workplaces described by Drury and Wick (1984) could be evaluated to show changes in productivity and reductions in cumulative trauma disorder causal factors. Similarly, Hasslequist (1981) showed productivity, quality, safety, and job satisfaction following human factors interventions in a computer component assembly line. In both cases, the objectives of the intervention were used to establish appropriate measures for the evaluation.

Ergonomics/human factors, however, is no longer confined to operating in a project mode. Increasingly, the establishment of a permanent function within an industry has meant that ergonomics is more closely related to the strategic objectives of the company. As Drury et al. (1989) have observed, this development requires measurement methodologies that also operate at the strategic level. For example, as a human factors group becomes more involved in strategic decisions about identifying and choosing the projects it performs, evaluation of the individual projects is less revealing. All projects performed could have a positive impact, but the group could still have achieved more with a more astute choice of projects. It could conceivably have had a more beneficial impact on the company's strategic objectives by stopping all projects for a period to concentrate on training the management, workforce, and engineering staff to make more use of ergonomics.

Such changes in the structure of the ergonomics/human factors profession indeed demand different evaluation methodologies. A powerful network of individuals, for example, who can, and do, call for human factors input in a timely manner can help an enterprise more than a number of individually successful project outcomes. Audit programs are one of the ways in which such evaluations can be made, allowing a company to focus its human factors resources most effectively. They can also be used in a prospective, rather than retrospective, manner to help quantify the needs of the company for ergonomics/human factors. Finally, they can be used to determine which divisions, plants, departments, or even product lines are in most need of ergonomics input.

5.2 Design Requirements for Audit Systems

Returning to the definition of an audit, the emphasis is on checking, acceptable policies, and consistency. The aim is to provide a fair representation of the business for use by third parties. A typical audit by a certified public accountant follows the steps outlined in the previous section (diagnostic investigation, transaction test, balances test, opinion formation).

Such a procedure can also form a logical basis for human factors audits. The first step chooses the areas of study, the second samples the system, the third analyzes these samples, and the final step produces an audit report. These define the broad issues in human factors audit design:

1. *How to sample the system.* How many samples are to be used, and how are they distributed across the system?
2. *What to sample.* What specific factors are to be measured, from biomechanical to organizational?
3. *How to evaluate the sample.* What standards, good practices, or ergonomic principles are to be used for comparison?
4. *How to communicate the results.* What techniques are to be used for summarizing the findings, and how far can separate findings be combined?

A suitable audit system needs to address all of these issues, but some overriding design requirements must first be specified.

5.2.1 Breadth, Depth, and Application Time

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. Thus, a thermal audit (Parsons, 1992) sacrifices breadth to provide considerable depth based on the heat balance equation but requires measurement of seven variables. Some can be obtained rapidly (air temperature, relative humidity), but some take longer (clothing insulation value, metabolic rate). Conversely, structured interviews with participants in an ergonomics program (Drury, 1990a) can be broad and rapid but quite deficient in depth.

At the level of audit instruments such as questionnaires or checklists, there are comprehensive surveys such as the Position Analysis Questionnaire (McCormick, 1979); the Arbeitswissenschaftliche Erhebungsverfahren zur Tätikgkeitsanalyse (AET) (Rohmert and Landau, 1989), which takes 2–3 h to complete; or the simpler Work Analysis Checklist (Pulat, 1992). Alternatively, there are simple single-page checklists such as the Ergonomics-Working Position-Sitting Checklist (SHARE, 1990), which can be completed in a few minutes. Analysis and reporting can range in depth from merely tabulating the number of ergonomic standards violated to expert systems that provide prescriptive interventions (Ayoub and Mital, 1989).

Most methodologies fall between the various extremes given above, but the goal of an audit system with an optimum trade-off between breadth, depth, and time is probably not realizable. A better practical course would be to select several instruments and use them together to provide the specific breadth and depth required for a particular application.

5.2.2 Use of Standards

The human factors/ergonomics profession has many standards and good practice recommendations. These differ by country [American National Standards Institute (ANSI), British Standards Institution (BSI), German Institute for Standardization (DIN)], although commonality is increasing through joint standards such as those of the International Organization for Standardization (ISO). Some standards are quantitative, such as heights for school furniture (BSI, 1980), sizes of characters or a video terminal display (VDT) screen (ANSI/HFES-200), and occupational exposure to noise. Other standards are more general in nature, particularly those which involve management actions to prevent or alleviate problems, such as the Occupational Safety and Health Administration (OSHA, 1990) guidelines for meatpacking plants. Generally, standards are more likely to exist for simple tasks and environmental stressors and are hardly to be expected for the complex cognitive activities with which human factors predictions increasingly deal. Where standards exist, they can represent unequivocal elements of audit procedures as a workplace that does not meet these standards is in a position of legal violation. A human factors program that tolerates such legal exposure should clearly be held accountable in any audit. A comprehensive listing of standards pertaining to human factors and ergonomics can be found in the appropriate handbook (Karwowski, 2005).

Merely meeting legal requirements, however, is an insufficient test of the quality of ergonomics/human factors efforts. Many legal requirements are arbitrary or outdated: for example, weight limits for manual materials handing in some countries. Additionally, other aspects of a job with high ergonomic importance may not be covered by standards: for example, the presence of multiple stressors, work in restricted spaces resulting in awkward postures, or highly repetitive upper extremity motions. Finally, there are many "human factors good practices" that are not the subject of legal standards. Examples are the National Institute for Occupational Safety and Health (NIOSH) lifting equation (Waters et al., 1993), the Illuminating Engineering Society (IES, 1993) codes, or the zones of thermal comfort defined by the American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE, 1989) or Fanger (1970). In some cases, standards are available in a different jurisdiction from that being audited. As an example, the military standard MIL-1472D (U.S. DoD, 1989) provides detailed standards for control and display design that are equally appropriate to process controls in manufacturing industry but have no legal weight there.

Despite the lack of legislation covering many human factors concerns, standards and other instantiations of good practice do have a place in ergonomics audits. Where they exist, they can be incorporated into an audit system without becoming the only criterion. Thus, noise levels in the United States have a legal limit of 90 dBA for hearing protection purposes. But at levels far below this, noise can disrupt communications (Jones and Broadbent, 1987) and distract from task performance. An audit procedure can assess the noise on multiple criteria (i.e., on hearing protection and on communication interruptions), with the former criterion used on all jobs and the latter only where verbal communication is an issue.

If standards and other good practices are used in a human factors audit, they provide a quantitative basis for decision making. Measurement reliability can be high and validity self-evident for legal standards. However, it is good practice in auditing to record only the measurement used, and not its relationship to the standard, which can be established later. This removes any temptation by the analyst to "bend" the measurement to reach a predetermined conclusion or to become complacent when the measurement is somewhat below the standard yet still potentially a detriment to human performance. Illumination measurements, for example, can vary considerably over a workspace, so that the audit question.

Is the work surface illumination > 750 lux?

☐ yes ☐ no

could be answered legitimately either way for some workspaces by choice of sampling point. Such temptation can be removed, for example, by the following audit question:

What is the illumination at four points on the workstation?

☐ ☐ ☐ ☐ lux

Later analysis can establish whether, for example, the mean exceeds 750 lux or whether any of the four points fall below this level.

It is also possible to provide later analyses that combine the effects of several simple checklist responses, as in Parsons's (1992) thermal audit, where no single measure would exceed good practice even though the overall result would be cumulative heat stress.

5.2.3 Evaluation of an Audit System

For a methodology to be of value, it must demonstrate validity, reliability, sensitivity, and usability. Most texts that cover measurement theory treat these aspects in detail (e.g., Kerlinger, 1964). Shorter treatments are found within human factors methodology texts (e.g., Drury, 1990b; Osburn, 1987).

Validity is the extent to which a methodology measures the phenomenon of interest. Does our ergonomics audit program indeed measure the quality of ergonomics in the plant? It is possible to measure validity in a number of ways, but ultimately all are open to argument. For example, if we do not know the "true" value of the "quality of ergonomics" in a plant, how can we validate our ergonomics audit program? Broadly, there are three ways in which validation can be tested.

Content validity is perhaps the simplest but least convincing measure. If each item of our measurement device displays the correct content, validity is established. Theoretically, if we could list all of the possible measures of a phenomenon, content validity would describe how well our measurement device samples these possible measures. In practice, it is assessed by having experts in the field judge each item for how well its content represents the phenomenon studied. Thus, the heat balance equation would be judged by most thermal physiologists to have a content that well represents the thermal load on an operator. Not all aspects are as easily validated!

A recent investigation of the content validity of five occupational health and safety management audits (Robson et al., 2010) against a published Canadian standard on occupational safety and health management found that 74% of the standard's content was partially (40%) or fully (34%) represented across all audits. In some cases, particular management elements audited were not covered well with considerable variability across program elements. Thus, even seemingly straightforward content validity of following a prescriptive standard may not be realized in practice if careful attention is not paid during the development and testing of the audit prior to implementation.

Concurrent (or *prediction*) *validity* has the most immediate practical impact. It measures empirically how well the output of the measurement device correlates with the phenomenon of interest. Of course, we must have an independent measure of the phenomenon of interest, which raises difficulties. To continue our example, if we used the heat balance equation to assess the thermal load on operators, there should be a high correlation between this and other measures of the effects of thermal load. Perhaps measures such as frequency of temperature complaints or of heat disorders: heat stroke, hyperthermia, hypothermia, and so on. In practice, however, measuring such correlations would be contaminated by, for example, the propensity to report temperature problems or individual acclimatization to heat. Overall outputs from a human factors audit (if such overall outputs have any useful meaning) should correlate with other measures of ergonomic inadequacy, such as injuries, turnover, quality measures, or productivity. Alternatively, we can ask how well the audit findings agree with independent assessments of qualified human factors engineers (Keyserling et al., 1992; Koli et al., 1993) and thus validate against one interpretation of current good practice.

Finally, there is *construct validity*. This is concerned with inferences made from scores, evaluated by considering all empirical evidence and models. Thus, a model may predict that one of the variables being measured should have a particular relationship to another variable not in the measurement device. Confirming this relationship empirically would help validate the particular construct underlying our measured variable. Note that different parts of an overall measurement device can have their construct validity tested in different ways. Thus, in a broad human factors audit, the thermal load could differentiate between groups of operators who do and do not suffer from thermal complaints. In the same audit a measure of difficulty in a target aiming task could be validated against Fitts's law. Other ways to assess construct validity are those that analyze clusters or factors within a group of measures. Different workplaces audited on a variety of measures and the scores, which are then subjected to factor analysis, should show an interpretable, logical structure in the factors derived. This method has been used on large databases for job-evaluation-oriented systems such as McCormick's position analysis questionnaire (PAQ) (McCormick, 1979).

Reliability refers to how well a measurement device can repeat a measurement on the same sample unit. Classically, if a measurement X is assumed to be composed of a true value X_t and a random measurement error X_e, then

$$X = X_t + X_e$$

For uncorrelated X_t and X_e, taking variances gives

$$\text{Variance}(X) = \text{Variance}(X_t) + \text{Variance}(X_e)$$

or

$$\text{V}(X) = \text{V}(X_t) + \text{V}(X_e)$$

We can define the reliability of the measurement as the fraction of measurement variance accounted for by true measurement variance:

$$\text{reliability} = \frac{\text{V}(X_t)}{\text{V}(X_t) + \text{V}(X_e)}$$

Typically, reliability is measured by correlating the scores obtained through repeated measurements. In an audit instrument, this is often done by having two (or more) auditors use the instrument on the same set of workplaces. The square of the correlation coefficient between the scores (either overall scores or separately for each logical construct) is then the reliability. Thus, PAQ was found to have an overall reliability of 0.79, tested using 62 jobs and two trained analysts (McCormick, 1979).

In the absence of direct measurement capabilities, many checklists rely on observation of workers and subsequent categorization (i.e., rating) of parameters of interest. Postures are an excellent example since most checklists utilize posture categories rather than relying on the rater to estimate angles. This introduces the issues of inter- and intrarater reliability. Inter- and intrarater reliability can be used during the development process to optimize reliability by revising categories to achieve higher reliability. Assessments of interrater reliability are more common, exemplified by the recent investigation of Park et al. (2009) examining interrater reliabilities for postures, materials handling, and noise and vibration exposure among hospital employees. Village et al. (2009) reported interrater reliabilities (kappa values) ranging from 0.21 to 1.0 for categorization of gross postures, trunk angles, and several materials-handling parameters.

Sensitivity defines how well a measurement device differentiates between entities. Does an audit system for human–computer interaction find a difference between software generally acknowledged to be "good" and "bad"? If not, perhaps the audit system lacks sensitivity, although of course there may truly be no difference between the systems except blind prejudice. Sensitivity can be affected adversely by poor reliability, which increases the variability in a measurement relative to a fixed difference between entities (i.e., gives a poor signal-to-noise ratio). Low sensitivity can also come from a floor or ceiling effect. These arise where almost all of the measurements cluster at a high or low limit. For example, if an audit question on the visual environment was

Does illumination exceed 10 lux?

☐ yes ☐ no

almost all workplaces could answer "yes" (although the author has found a number that could not meet even this low criterion). Conversely, a floor effect would be a very high threshold for illuminance. Sensitivity can arise too when validity is in question. Thus, heart rate is a valid indicator of heat stress but not of cold stress. Hence, exposure to various degrees of cold stress would be measured only insensitively by heart rate.

Usability refers to the auditor's ease of use of the audit system. Good human factors principles should be followed: for example, document design guidelines in constructing checklists (Wright and Barnard, 1975; Patel et al., 1993). If the instrument does not have good usability, it will be used less often and may even show reduced reliability due to auditors' errors.

5.3 Audit System Design

As outlined in Section 2, the audit system must choose a sample, measure the sample, evaluate it, and communicate the results. In this section we approach these issues systematically.

An audit system is not just a checklist; it is a methodology that often includes the technique of a checklist. The distinction needs to be made between methodology and techniques. Almost three decades ago, Easterby (1967) used Bainbridge and Beishon's (1964) definitions:

- *Methodology*: a principle for defining the necessary procedures
- *Technique*: a means to execute a procedural step

Easterby notes that a technique may be applicable in more than one methodology.

5.3.1 Sampling Scheme

In any sampling, we must define the unit of sampling, the sampling frame, and the sample choice technique.

For a human factors audit the unit of sampling is not as self-evident as it appears. From a job evaluation viewpoint (e.g., McCormick, 1979), the natural unit is the job that is composed of a number of tasks. From a medical viewpoint the unit would be the individual. Human factors studies focus on the task/operator/machine/environment (TOME) system (Drury, 1992a,b) or, equivalently, the software/hardware/environment/liveware (SHEL) system [International Civil Aviation Organization (ICAO), 1989]. Thus, from a strictly human factors viewpoint, the specific combination of TOME can become the sampling unit for an audit program.

Unfortunately, this simple view does not cover all the situations for which an audit program may be needed. Although it works well for the rather repetitive tasks performed at a single workplace typical of much manufacturing and service industry, it cannot suffice when these conditions do not hold. One relaxation is to remove the stipulation of a particular incumbent, allowing for jobs that require frequent rotation of tasks. This means that the results for one task will depend on the incumbent chosen or that several tasks will need to be combined if an individual operator is of interest. A second relaxation is that the same operator may move to different workplaces, thus changing the environment as well as the task. This is typical of maintenance activities, where a mechanic may perform any one of a repertory of hundreds of tasks, rarely repeating the same task. Here, the rational sampling unit is the task, which is observed for a particular operator at a particular machine in a particular environment. Examples of audits of repetitive tasks (Mir, 1982; Drury, 1990a) and maintenance tasks (Chervak and Drury, 1995) are given later to illustrate these different approaches.

Definition of the sampling frame, once the sampling unit is settled, is more straightforward. Whether the frame covers a department, a plant, a division, or an entire company, enumeration of all sampling units is possible at least theoretically. All workplaces, or jobs, or individuals can in principle be listed, although in practice the list may never be up to date in an agile industry where change is the normal state of affairs. Individuals can be listed from personnel records, tasks from work orders or planning documents, and workplaces from plant layout plans. A greater challenge, perhaps, is to decide whether indeed the entire plant really is the focus of the audit. Do we include office jobs or just production? What about managers, foremen, part-time janitors, and so on? A good human factors program would see all of these tasks or people as worthy of study, but in practice they may have had different levels of ergonomic effort expended upon them. Should some tasks or groups be excluded from the audit merely because most participants agree that they have few pressing human factors problems? These are issues that need to be decided explicitly before the audit sampling begins.

Choice of the sample from the sampling frame is well covered in sociology texts. Within human factors it typically arises in the context of survey design (Sinclair, 1990). To make statistical inferences from the sample to the population (specifically to the sampling frame), our sampling procedure must allow the laws of probability to be applied. The sampling methods used most often are described here.

Random Sampling Each unit within the sampling frame is equally likely to be chosen for the sample. This is the simplest and most robust method, but it may not be the most efficient. Where subgroups of interest (strata) exist and these subgroups are not equally represented in the sampling frame, one collects unnecessary information on the most populous subgroups and insufficient information on the least populous. This is because our ability to estimate a population statistic from a sample depends on the absolute sample size and not, in most practical cases, on the population size. As a corollary, if subgroups are of no interest, random sampling loses nothing in efficiency.

Stratified Random Sampling Each unit within a particular stratum of the sampling frame is equally likely to be chosen for the sample. With stratified random sampling we can make valid inferences about each of the strata. By weighting the statistics to reflect the size of the strata within the sampling frame, we can also obtain population inferences. This is often the preferred auditing sampling method, as, for example, we would wish to distinguish between different classes of tasks in our audits: production, warehouse, office, management, maintenance, security, and so on. In this way our audit interpretation could give more useful information concerning where ergonomics is being used appropriately.

Cluster Sampling Clusters of units within the sampling frame are selected, followed by random or nonrandom selection within clusters. Examples of clusters would be the selection of particular production lines within a plant (Drury, 1990a) or selection of "representative" plants within a company or division. The difference between cluster and stratified sampling is that in cluster sampling only a subset of possible units within the sampling frame is selected, whereas in stratified sampling all of the sampling frame is used, as each unit must belong to one stratum. Because clusters are not randomly selected, the overall sample results will not reflect population values, so that statistical inference is not possible. If units are chosen randomly within each cluster, statistical inference within each cluster is possible. For example, if three production lines are chosen as clusters and workplaces sampled randomly within each, the clusters can be regarded as fixed levels of a factor and the data subjected to analysis of variance to determine whether there are significant differences between levels of that factor. What is sacrificed in cluster sampling is the ability to make *population* statements. Continuing this example, we could state that the lighting in line A is better than in line B or C but still not be able to make statistically valid statements about the plant as a whole.

5.3.2 Data Collection Instrument

So far we have assumed that the instrument used to collect the data from the sample is based on measured data where appropriate. Although this is true of many audit instruments, this is not the only way to collect audit data. There have been interviews with participants (Drury, 1990a), interviews and group meetings to locate potential errors (Fox, 1992), and use of archival data such as injury of quality records (Mir, 1982). All have potential uses with, as remarked earlier, a judicious range of methods often providing the appropriate composite audit system.

One consideration regarding audit technique design and use is the extent of computer involvement. Computers are now inexpensive, portable, and powerful, so that they can be used to assist data collection, data verification, data reduction, and data analysis (Drury, 1990a). With the advent of more intelligent interfaces, checklist questions can be answered from mouse clicks on buttons, or selection from menus, as well as the more usual keyboard entry. Data verification can take place at entry time by checking for out-of-limits data, or odd data, such as the ratio of luminance to illuminance implying a reflectivity greater than 100%. In addition, branching in checklists can be made easier, with only valid follow-on questions highlighted. The "checklist user's manual" can be built into the checklist software using context-sensitive help facilities, as in the ergonomics evaluation analysis methodology (EEAM) checklist (Chervak and Drury, 1995). Computers can, of course, be used for data reduction (e.g., finding the insulation value of clothing from a clothing inventory), data analysis, and results presentation.

Having made the case for computer use, some precautions are in order. Computers are still bulkier than simple pencil-and-paper checklists. Computer reliability is not perfect, so that inadvertent data loss is still a real possibility. Finally, software and hardware date much more rapidly than hard copy, so that results stored safely on the latest media may be unreadable 10 years later. How many of us can still read punched cards or 8-in. floppy disks? In contrast, hard-copy records are still available from before the start of the common era.

Checklists and Surveys as Audit Tools For many practitioners the proof of the effectiveness of an ergonomics effort lies in the ergonomic quality of the work systems it produces. A plant or office with appropriate human–machine function allocation, well-designed workplaces, comfortable environment, adequate placement/training, and inherently satisfying jobs almost by definition has been well served by human factors. Such a facility may not have human factors specialists, just good designers of environment, training, organization, and so on, working independently, but this would generally be a rare occurrence. Thus, a checklist to measure such inherently ergonomic qualities has great appeal as part of an audit system. We have covered the design aspects of checklists in general, so we concentrate here on their use in the context of human factors/ergonomics audits.

Such checklists are almost as old as the discipline. An early paper by Burger and deJong (1964) lists four earlier checklists for ergonomic job analysis before going on to develop their own. Theirs was commissioned by the International Ergonomics Association (IEA) in 1961 and is usually known as the IEA checklist. It was based in part on one developed at the Philips Health Centre by G. J. Fortuin and provided in detail in Burger and deJong's paper.

Like any other questionnaire, a checklist needs to have both a helpful overall structure and well-constructed questions. It should also be proven reliable, valid, sensitive, and usable, although precious few meet all these criteria. A recent survey of Certified Professional Ergonomists in the United States (Dempsey et al. 2005) revealed that 70.5% of respondents used checklists, with 67 of 301 respondents that reported using checklists indicating the the checklist was developed on their own or by their employer. It is unlikely these were subjected to tests of reliability or validity, although the usability should be high for those developed by the actual users. The most commonly identified single checklist was referred to as the OSHA checklist, with only 27 responses indicating that there are many checklists used by practicing ergonomists with seemingly few highly popular checklists.

In the remainder of this section, a selection of checklists is presented as typical of (reasonably) good practice. Emphasis will be on objective, structure, and question design. Note that checklists are not the only approach possible. Westwater and Johnson (1995) compared them with expert evaluation and empirical user testing in evaluating PDA design. They concluded that user-based evaluations led to more insights for this evaluation.

1. *IEA Checklist*. The IEA checklist (Burger and deJong, 1964) was designed for ergonomic job analysis over a wide range of jobs. It uses the concept of functional load to give a logical framework relating physical load, perceptual load, and mental load to the worker, the environment, and working methods/tools/machines. Within each cell (or subcell, e.g., physical load could be static or dynamic) the load was assessed on different criteria, such as force, time, distance, occupational medical, and psychological criteria. Table 2 shows the structure and typical questions. Dirken (1969) modified the IEA checklist to improve the questions and methods of recording. He found that it could be applied in a median time of 60 min per workstation. No data are given on evaluation of the IEA checklist, but its structure has been so influential that it is included here for more than historical interest.
2. *Position Analysis Questionnaire*. The PAQ is a structured job analysis questionnaire using 187 worker-oriented elements to characterize the human behaviors involved in jobs (McCormick et al., 1969). The PAQ is structured into six divisions, with the first three representing the classic experimental psychology approach (information input, mental process, work output) and the next a broader sociotechnical view (relationships with other persons, job context, other job characteristics). Table 3 shows these major divisions, examples of job elementsin each, and the rating scales employed for response.

Table 2 IEA Checklist Structure and Typical Questions

Structure

		A	B	C
Load: 1. Mean 2. Peaks (intensity, frequency, duration)		Worker	Environment	Working Method, Tools, Machines
I. Physical load	1. Dynamic			
	2. Static			
II. Perceptual load	1. Perception			
	2. Selection, decision			
	3. Control of movement			
III. Mental load	1. Individual			
	2. Group			

Typical Question

I/B. Physical Load/Environment	2.1. Physiological Criteria
	1. Climate: high and low temperatures 1. Are these extreme enough to affect comfort or efficiency? 2. If so, is there any remedy? 3. To what extent is working capacity adversely affected? 4. Do personnel have to be specially selected for work in this particular environment?

Table 3 PAQ Structure and Scales

Structure

Division	Definition	Examples of Questions
1. Information input	Where and how does the worker get the information he uses in performing his job?	1. Use of written materials 2. Near-visual differentiation
2. Mental processes	What reasoning, decision making, planning, and information-processing activities are involved in performing the job?	1. Level of reasoning in problem solving 2. Coding–decoding
3. Work output	What physical activities does the worker perform, and what tools or devices does he use?	1. Use of keyboard devices 2. Assembling–unassembling
4. Relationships with other persons	What relationships with other people are required in performing the job?	1. Instructing 2. Contacts with public or customers
5. Job context	In what physical or social contexts is the work performed?	1. High temperature 2. Interpersonal; conflict situations
6. Other job characteristics	What activities, conditions, or characteristics other than those described above are relevant to the job?	1. Specified work pace 2. Amount of job structure

Scales

Types of Scales		Scales Values	
Code	Type of Rating	Rating	Definition
U	Extent of use	N	Does not apply
I	Importance of the job	1	Very minor
T	Amount of time	2	Low
P	Possibility of occurrence	3	Average
A	Applicability (yes/no only)	4	High
S	Special code	5	Extreme

Source: McCormick (1979).

Construct validity was tested by factor analyses of databases containing 3700 and 2200 jobs, which established 45 factors. Thirty-two of these fit neatly into the original six-division framework, with the remaining 13 being classified as "overall dimensions." Further proof of construct validity was based on 76 human attributes derived from the PAQ, rated by industrial psychologists and the ratings subjected to principal-component analysis to develop dimensions "which had reasonably similar attribute profiles" (McCormick, 1979, p. 204). As noted earlier, interreliability was 0.79 based on another sample of 62 jobs.

The PAQ covers many of the elements of concern to human factors engineers and has indeed much influenced subsequent instruments, such as AET. With good reliability and useful (although perhaps dated) construct validity, it is still a useful instrument if the natural unit of sampling is the job. The exclusive reliance on rating scales applied by the analyst goes rather against current practice of comparison of measurements against standards or good practices.

3. *AET* (*Arbeitswissenschaftliche Erhebungsverfahren zur Tätikgkeitsanalyse*). The AET has been published in German (Landau and Rohmert, 1981) and later in English (Rohmert and Landau, 1983). It is the job analysis subsystem of a comprehensive system of work studies. It covers "the analysis of individual components of man-at-work systems as well as the description and scaling of their interdependencies" (Rohmert and Landau, 1983, pp. 9–10). As with all good techniques, it starts from a model of the system (Verband für Arbeitsgestaltung, Betriebsorganisation und Unternehmensentwicklung, REFA, 1971; referenced in Wagner, 1989), to which is added Rohmert's stress–strain concept. The latter sees strain as being caused by the intensity and duration of stresses impinging on an operator's individual characteristics. It is seen as useful in the analysis of requirements and work design, organization in industry, personnel management, and vocational counseling and research.

AET itself was developed over many years using PAQ as an initial starting point. Table 4 shows the structure of the survey instrument with typical questions and rating scales. Note the similarity between AET's job demands analysis and the first three categories of the PAQ and between the scales used in AET and PAQ (Table 3).

Measurements of validity and reliability of AET are discussed by H. Luczak in an appendix to Landau and Rohmert (1981), although no numerical values are given. Cluster analysis of 99 AET records produced groupings that supported the AET constructs. Seeber et al. (1989) used AET along with two other work analysis methods in 170 workplaces. They found that AET provided the most differentiating aspects (suggesting sensitivity). They also measured postural complaints and showed that only the AET groupings for 152 female workers found significant differences between complaint levels, thus helping establish construct validity.

Like PAQ before it, AET has been used on many thousands of jobs, mainly in Europe. A sizable database is maintained that can be used for both norming of new jobs analyzed and analysis to test research hypotheses. It remains a most useful instrument for work analysis.

4. *Ergonomics Audit Program* (Mir, 1982; Drury, 1990a). This program was developed at the request of a multinational corporation to be able to audit its various divisions and plants as ergonomics programs were being instituted. The system developed was a methodology of which the workplace survey was one technique. Overall, the methodology used archival data or outcome measures (injury reports, personnel records, productivity) and critical incidents to rank order departments within a plant. A cluster sampling of these departments gives either the ones with the highest need (if the aim is to focus ergonomic effort) or a sample representative of the plant (if the objective is an audit). The workplace survey is then performed on the sampled departments.

The workplace survey was designed based on ergonomic aspects derived from a task/operator/machine/environment model of the person at work. Each aspect formed a section of the audit, and sections could be omitted if they were clearly not relevant (e.g., manual materials-handling aspects for data entry clerks). Questions within each section were based on standards, guidelines, and models, such as the NIOSH (1981) lifting equation, ASHRAE's (1990) *Handbook of Fundamentals* for thermal aspects, and Givoni and Goldman's (1972) model for predicting heart rate. Table 5 shows the major sections and typical questions.

Data were entered into the computer program and a rule-based logic evaluated each section to provide messages to the user in the form of either a "section shows no ergonomic problems" message:

MESSAGE
Results from analysis of auditory aspects:
Everything OK in this section.

or discrepancies from a single input:

MESSAGE
Seats should be padded, covered with nonslip materials, and have the front edge rounded.

Table 4 AET Structure and Scales

Structure

Part	Major Divisions	Sections
A. Work systems analysis	1. Work objects	1.1. Material work objects 1.2. Energy as work object 1.3. Information as work object 1.4. Humans, animals; plants as work objects
	2. Equipment	2.1. Working equipment 2.2. Other equipment
	3. Work environment	3.1. Physical environment 3.2. Organizational and social environment 3.3. Principles and methods of remuneration
B. Task analysis	1. Tasks relating to material work objects 2. Tasks relating to abstract work objects 3. Human-related tasks 4. Number and repetitiveness of tasks	
C. Job demand analysis	1. Demands on perception	1.1. Mode of perception 1.2. Absolute/relative evaluation of perceived information 1.3. Accuracy of perception
	2. Demands for decision	2.1. Complexity of decisions 2.2. Pressure of time 2.3. Required knowledge
	3. Demands for response/activity	3.1. Body postures 3.2. Static work 3.3. Heavy muscular work 3.4. Light muscular work, active light work 3.5. Strenuousness and frequency of movements

Scales

Types of Scales		Scales Values	
Code	Type of Rating	Duration Value	Definition
A	Does this apply?	0	Very infrequent
F	Frequency	1	Less than 10% of shift time
S	Significance	2	Less than 30% of shift time
D	Duration	3	30–60% of shift time
		4	More than 60% of shift time
		5	Almost continuously during whole shift

or discrepancies based on the integration of several inputs:

MESSAGE

The total metabolic workload is 174 watts.

Intrinsic clothing insulation is 0.56 clo.

Initial rectal temperature is predicted to be 36.0°C.

Final rectal temperature is predicted to be 37.1°C.

Counts of discrepancies were used to evaluate departments by ergonomics aspect, while the messages were used to alert company personnel to potential design changes. The latter use of the output as a training device for nonergonomic personnel was seen as desirable in a multinational company rapidly expanding its ergonomics program.

Reliability and validity have not been assessed, although the checklist has been used in a number of industries (Drury, 1990a). The workplace survey has been included here because, despite its lack of measured reliability and validity, it shows the relationship between audit as methodology and checklist as technique.

5. *ERGO, EEAM, and ERNAP* (Koli et al., 1993; Chervak and Drury, 1995). These checklists are both part of complete audit systems for different aspects of civil aircraft hangar activities. They were developed for the FAA to provide tools for assessing human factors in aircraft inspection (ERGO) and maintenance (EEAM)

Table 5 Workplace Survey Structure and Typical Questions

Section	Major Classification	Examples of Questions
1. Visual aspects		Nature of task? Illuminance at task (midfield, outer field)?
2. Auditory aspects		Noise level (dBA)? Main source of noise?
3. Thermal aspects		Strong radiant sources present? Wet bulb temperature? (Clothing inventory)
4. Instruments, controls, displays	Standing vs. seated	Are controls mounted between 30 and 70 in.?
	Displays	Signals for crucial visual checks?
	Labeling	Are trade names deleted?
	Coding	Color codes same for control and display?
	Scales, dials, counters	All numbers upright on fixed scales?
	Control–display relationships	Grouping by sequence or subsystem?
	Controls	Emergency button diameter > 0.75 in.?
5. Design of workplaces	Desks	Seat to underside of desk > 6.7 in.?
	Chairs	Height easily adjustable to 15–21 in.?
	Posture	Upper arms vertical?
6. Manual materials handling	NIOSH (1981) lifting guide	Task, H, V, D, F
7. Energy expenditure		Cycle time? Object weight? Type of work?
8. Assembly/ repetitive aspects		Seated, standing, or both? If heavy work, is bench 6–16 in. below elbow height?
9. Inspection aspects		Number of fault types? Training time until unsupervised?

activities, respectively. Inspection and maintenance activities are nonrepetitive in nature, controlled by task cards issued to technicians at the start of each shift. Thus, the sampling unit is the task card, not the workplace, which is highly variable between task cards. Their structure was based on extensive task analyses of inspection and maintenance tasks, which led to generic function descriptions of both types of work (Drury et al., 1990). Both systems have sampling schemes and checklists. Both are computer based with initial data collection on either hard copy or direct into a portable computer. Recently, both have been combined into a single program (ERNAP) distributed by the FAA's Office of Aviation Medicine. The structure of ERNAP and typical questions are given in Table 6.

As in Mir's ergonomics audit program, the ERNAP, the checklist is again modular, and the software allows formation of data files, selection of required modules, analysis after data entry is completed, and printing of audit reports. Similarly, the ERGO, EEAM, and ERNAP instruments use quantitative or yes/no questions comparing the value entered with standards and good-practice guides. Each takes about 30 min per task. Output is in the form of an audit report for each workplace, similar to the messages given by Mir's workplace survey, but in narrative form. Output in this form was chosen for compatibility with existing performance and compliance audits used by the aviation maintenance community.

Reliability of a first version of ERGO was measured by comparing the output of two auditors on three tasks. Significant differences were found at $p < 0.05$ on all three tasks, showing a lack of interrater reliability. Analysis of these differences showed them to be due largely to errors on questions requiring auditor judgment. When such questions were replaced with more quantitative questions, the two auditors had no significant disagreements on a later test. Validity was measured using concurrent validation against six Ph.D. human factors engineers who were asked to list all ergonomic issues on a power plant inspection task. The checklist found more ergonomic issues than the human factors engineers. Only a small number of issues were raised by the engineers that were missed by the checklist. For the EEAM checklist, again an initial version was tested for reliability with two auditors and achieved the same outcome for only 85% of the questions. A modified version was tested and the reliability was considered satisfactory with 93% agreement. Validity was again tested against four human factors engineers;

Table 6 ERNAP Structure and Typical Questions

Audit Phase	Major Classification	Examples of Questions
I. Premaintenance	Documentation	Is feedforward information on faults given?
	Communication	Is shift change documented?
	Visual characteristics	If fluorescent bulbs are used, does flicker exist?
	Electric/pneumatic equipment	Do pushbuttons prevent slipping of fingers?
	Access equipment	Do ladders lave nonskid surfaces on landings?
II. Maintenance	Documentation	Does inspector sign off workcard after each task?
	Communication	Explicit verbal instructions from supervisor?
	Task lighting	Light levels in four zones during task (fc)?
	Thermal issues	Wet-bulb temperature in hanger bay (°C)?
	Operator perception	Satisfied with summer thermal environment?
	Auditory issues	Noise levels at five times during task (dBA)?
	Electrical and pneumatic	Are controls easily differentiated by touch?
	Access equipment	Is correct access equipment available?
	Hand tools	Does the tool handle end in the palm?
	Force measurements	What force is being applied (kg)?
	Manual material handling	Does task require pushing or pulling forces?
	Vibration	What is total duration of exposure on this shift?
	Repetitive motion	Does the task require flexion of the wrist?
	Access	How often was access equipment repositioned?
	Posture	How often were following postures adopted?
	Safety	Is inspection area cleaned adequately for inspection?
	Hazardous material	Were hazardous materials signed out and in?
III. Postmaintenance	Buyback	Are discrepancy worksheets readable?

this time the checklist found significantly more ergonomic issues than the engineers without missing any of the issues they raised.

The ERNAP audits have been included here to provide examples of a checklist embedded in an audit system where the workplace is *not* the sampling unit. They show that nonrepetitive tasks can be audited in a valid and reliable manner. In addition, they demonstrate how domain-specific audits can be designed to take advantage of human factors analyses already made in the domain.

6. *Upper Extremity Checklist* (Keyserling et al., 1993). As its name suggests, this checklist is narrowly focused on biomechanical stresses to the upper extremities that could lead to cumulative trauma disorders (CTDs). It does not claim to be a full-spectrum analysis tool but is included here as a good example of a special-purpose checklist that has been carefully constructed and validated. The checklist (Table 7) was designed for use by management and labor to fulfill a requirement in the OSHA guidelines for meatpacking plants. The aim is to screen jobs rapidly for harmful exposures rather than to provide a diagnostic tool. Questions were designed based on the biomechanical literature, structured into six sections. Scoring was based on simple presence or absence of a condition or on a three-level duration score. As shown in Table 7, the two or three levels were scored as 0, √, or *, depending on the stress rating built into the questionnaire. These symbols represented insignificant, moderate, or substantial exposures. A total score could be obtained by summing moderate and substantial exposures.

The upper extremity checklist was designed to be biased toward false positives (i.e., to be very sensitive). It was validated against detailed analyses of 51 jobs by an ergonomics expert. Each section (except the first, which recorded only the dominant hand) was considered as giving a positive screening if at least one * rating was recorded. Across the various sections, there was reasonable agreement between checklist users and the expert analysis, with the checklist being generally more sensitive, as was its aim. The original reference shows the findings of the checklist applied to 335 manufacturing and warehouse jobs.

As a special-purpose technique in an area of high current visibility for human factors, the upper extremity checklist has proven validity, can be used by those with minimal ergonomics training for screening jobs, and takes only a few minutes per workstation. The same team has also developed and validated a legs, trunk, and neck job screening procedure along similar lines (Keyserling et al., 1992).

Table 7 Upper Extremity Checklist: Structure and Scoring

Structure

Major Section	Examples of Questions
Worker information	Which hand is dominant?
	Repetitive use of the hands and wrists?
	If "yes," then:
	Is cycle < 30 s?
	Repeated for >50% cycle?
Mechanical stress	Do hard or sharp objects put localized pressure on:
	Back or side of fingers?
	Palm or base of hand?
Force	Lift, carry, push, or pull objects >4.5 kg?
	If gloves worn, do they hinder gripping?
Posture	Is pinch grip used?
	Is there wrist deviation?
Tools, hand-held objects and equipment	Is vibration transmitted to the operator's hand?
	Does cold exhaust air blow on the hand or wristz?

Scoring Scheme

Question	Scoring		
Is there wrist deviation?	No	Some	> 33% cycle
	0	√	*
Overall evaluation:			
total score = number of √ + number of *			

7. *Ergonomic Checkpoints.* The workplace improvement in small enterprises (WISE) methodology (Kogi, 1994) was developed by the IEA and the International Labour Office (ILO) to provide cost-effective solutions for smaller organizations. It consists of a training program and a checklist of potential low-cost improvements. This checklist, called *ergonomics checkpoints*, can be used both as an aid to discovery of solutions and as an audit tool for workplaces within an enterprise.

   The 128-point checklist has now been published (Kogi and Kuorinka, 1995) and as a book by the ILO (1999). It covers the nine areas shown in Table 8. Each item is a statement rather than a question and is called a *checkpoint*. For each checkpoint there are four sections, also shown in Table 8. There is no scoring system as such; rather, each checkpoint becomes a point of evaluation of each workplace for which it is appropriate. Note that each checkpoint also covers why that improvement is important and a description of the core issues underlying it. Both of these help the move from rule-based reasoning to knowledge-based reasoning as nonergonomists continue to use the checklist. A similar idea was embodied in the Mir (1982) ergonomic checklist.
8. *Operational Demand Evaluation Checklist*. Public et al. (2010) developed this checklist to capture the operational demands on rail signalers. Rail signalers are responsible for directing rail traffic to attain timely performance as well as safety of the traffic as well as equipment and people that could be impacted by a mishap. Rail systems are complex dynamic

Table 8 Ergonomic Checkpoints

Structure of the Checklist

Major Section	Typical Checkpoints
Materials handling	Clear and mark transport ways.
Handtools	Provide handholds, grips, or good holding points for all packages and containers.
Productive machine safety	Use jigs and fixtures to make machine operations stable, safe, and efficient.
Improving workstation design	Adjust working height for each worker at elbow level or slightly below it.
Lighting	Provide local lights for precision or inspection work.
Premises	Ensure safe wiring connections for equipment and lights.
Control of hazards	Use feeding and ejection devices to keep hands away from dangerous parts of machinery.
Welfare facilities	Provide and maintain good changing, washing, and sanitary facilities to keep good hygiene and tidiness.
Work organization	Inform workers frequently about the results of their work.

Structure of Each Checkpoint

Why?	Reasons why improvements are important
How?	Description of several actions each of which can contribute to improvement
Some more hints	Additional points which are useful for attaining the improvement
Points to remember	Brief description of the core element of the checkpoint

Source: K. Kogi, private communication, November 13, 1995.

entities, and maintaining mental workload below levels that could lead to errors or inappropriate decisions is important.

Public et al. (2010) began the process of creating the checklist with extensive field data collection with signalers using a range of data collection methods, including interviews, observations, and verbal protocol analysis. A structured interview technique (repertory grid technique) was used to elicit knowledge from rail signalers about the most significant elements identified during initial data collection with respect to influencing signaler mental workload. The identified elements were grouped into categories of operational infrastructure, indicators, process, and service pattern. The repertory grid technique was used to understand which constructs were most meaningful to subject matter experts in describing the overall construct of mental workload and how the elements identified earlier were most able to reflect the presence or absence of the constructs.

The results of the repertory grid technique were then used to develop the checklist. The scoring of high, medium, and low for each element was determined from previously collected data or based upon subject matter experts' opinions where data were lacking. Concurrent validity was assessed by comparing the results of the checklist to a "grading system familiar to the rail industry." The comparison tool graded signal boxes on the "responsibilities and decision-making required as a consequence of the infrastructure controlled and the service operated." Interrater reliability was assessed, but no statistics were presented. Training was developed in an attempt to increase interrater reliability.

9. *Line-Oriented Safety Audit* (*LOSA*). Over 15 years the civil aviation industry has developed a safety audit program for cockpit crew. This was initially developed as a way to check whether the discipline of crew resource management (CRM: e.g., Helmreich et al., 2001) had taken hold in cockpit crews during line operations (Klinect et al., 2003). This was an in-cockpit observation method using trained observers (human factors engineers or experienced airline pilots). The methodology was later refined as a result of new research on threat error management (TEM) and is now an FAA Advisory Circular (FAA, 2006). The actual data collection methodology is quite different from the typical checklist, although it does have boxes to fill in. These are narrative boxes for each stage of flight: predeparture/ axi, take-off/climb, cruise, descent/approach/land/taxi. These raw data are verified by a data verification round table, which can take several days. From this verification comes a set of errors in each of several standard categories, finally being presented stripped of all identification material as error rates in each category. A formal reliability test of LOSA was conducted using the LOSA observer feedback form (LOFF) with two coding exercises using 116 trained observers. The Kudler–Richardson KR-20 coefficient was measured as 0.70, while the split-half Spearman–Brown coefficient was 0.88, both indications of high reliability (Klinect, 2005, pp. 70–71). Currently the LOSA is being extended to airline ground, ramp, and maintenance operations (Ma et al., 2011).
10. Smith (2001) described a production line audit designed to establish a baseline of current ergonomics activities and priorities. The approach used observational task analysis in conjunction with questionnaires and interviews. Significant musculoskeletal stressors were selected as the focus of the audit program, and thus this audit was not designed to audit the potential broad spectrum of ergonomics issues that could be present in an assembly environment. The audit relied on observational analyses of risk factors for each body part, supplemented by body part discomfort maps filled out by line employees, although no details of reliability or validity are given.
11. *Other Checklists*. The sample of successful audit checklists above has been presented in some detail to provide the reader with their philosophy, structure, and sample questions. Rather than continue in the same vein, other interesting checklists are outlined in Table 9. Each entry shows the domain, the types of issues addressed, the size or time taken in use, and whether validity and reliability have been measured. Most textbooks now provide checklists, and a few of these are cited. No claim is made that Table 9 is comprehensive; rather, it is a sampling with references so that readers can find a suitable match to their needs. The first nine entries in the table are conveniently colocated in Landau and Rohmert (1989). Many of their reliability and validity studies are reported in this publication. The next entries are results of the Commission of European Communities fifth ECSC program, reported in Berchem-Simon (1993). Others are from texts and original references. The author has not personally used all of these checklists and so cannot endorse them specifically. Also, omission of a checklist from this table implies nothing about its usefulness.

Other Data Collection Methods. Not all data come from checklists and questionnaires. We can audit a human factors program using outcome measures alone. However, outcome measures such as injuries, quality, and productivity are nonspecific to human factors: Many other external variables can affect them. An obvious example is changes in the reporting threshold for injuries, which can lead to sudden apparent increases and decreases in the safety of a department or plant.

Table 9 Selection of Published Checklists

Name	Reference	Coverage	Reliability	Validity
TBS	Hacker et al. (1983)	Mainly mental work		Vs. AET
VERA	Volpert et al. (1983)	Mainly mental work		Vs. AET
RNUR	RNUR (1976)	Mainly physical work		
LEST	Guélaud (1975)	Mainly physical work		
AVISEM	AVISEM (1977)	Mainly physical work		
GESIM	GESIM (1988)	Mainly physical work		
RHIA	Leitner and Greiner (1989)	Task hindrances, stress	0.53–0.79	Vs. many
MAS	Groth (1989)	Open structure, derived from AET		Vs. AET
JL and HA	Mattila and Kivi (1989)	Mental, physical work, hazards	0.87–0.95	
	Bolijn (1993)	Physical work for women	Tested	
	Panter (1993)	Load handling		
	Portillo Sosa (1993)	VDT standards		
Work analysis	Pulat (1992)	Mental and physical work		
Thermal audit	Parsons (1992)	Thermal audit from heat balance		Content
WAS	Yoshida and Ogawa (1991)	Workplace and environment	Tested	Vs. expert
Ergonomics	SHARE (1990)	Short workplace checklists		
	Cakir et al. (1980)	VDT checklist		
	Nery (1999)	Meat processing		
	Robson and Wolstenhulme (2010)	Ultrasound scan room		

Source: First nine from Landau and Rohmert (1989), next three from Berchem-Simon (1993)

Additionally, injuries are (or should be) extremely rare events. Thus, to obtain enough data to perform meaningful statistical analysis may require aggregation over many disparate locations andz/or time periods. In ergonomics audits, such outcome measures are perhaps best left for long-term validation or for use in selecting cluster samples. An example is the "validation" of LOSA by using it to measure the improvements in error reduction after a CRM program was instituted. Ma et al. (2010) quote Croft (2001) showing a 59% decline in unstabilized approaches as measured by LOSA after a CRM program at Continental Airlines.

Besides outcome measures, interviews represent a possible data collection method. Whether directed or not (e.g., Sinclair, 1990), they can produce critical incidents, human factors examples, or networks of communication (e.g., Drury, 1990a) that have value as part of an audit procedure. Interviews are used routinely as part of design audit procedures in large-scale operations such as nuclear power plants (Kirwan, 1989) or naval systems (Malone et al., 1988).

A novel interview-based audit system was proposed by Fox (1992) based on methods developed by British Coal (reported by Simpson, 1994). Here an error-based approach was taken using interviews and archival records to obtain a sampling of actual and possible errors. These were then classified using Reason's (1990) active/latent failure scheme and orthogonally by Rasmussen's (1987) skill-, rule-, and knowledge-based framework. Each active error is thus a conjunction of skill/mistake/violation with skill/rule/knowledge. Within each conjunction, performance-shaping factors can be deduced and sources of management intervention listed. This methodology has been used in a number of mining-related studies (see Section 5.4.2).

It is worth mentioning that the term ergonomics audit is occasionally used by consultants in reference to assessing elements of ergonomics programs either corporatewide or at individual sites. The audits do not necessarily assess the ergonomics of tasks, machines/equipment, or environment but rather assess whether ergonomics processes such as individual workstations are carried out, whether surveillance for injuries is carried out, the nature of policies and procedures, and so on. These are typically carried out through interviews. Although these may be effective for increasing the quality of an ergonomics program, they do not necessarily measure effectiveness of the program.

5.3.3 Data Analysis and Presentation

Human factors as a discipline covers a wide range of topics from workbench height to function allocation in automated systems. An audit program can only hope to abstract and present a part of this range. With our consideration of sampling systems and data collection devices we have seen different ways in which an unbiased abstraction can be aided. At this stage the data consist of large numbers of responses to large numbers of checklist items or detailed interview findings. How can, or should, these data be treated for best interpretation?

Here there are two opposing viewpoints: One is that the data are best summarized across sample units but not across topics. This is typically the way the human factors professional community treats the data, giving summaries in published papers of the distribution of responses to individual items on the checklist. In this way, findings can be more explicit: for example, that the lighting is an area that needs ergonomics effort or that the seating is generally poor. Adding together lighting and seating discrepancies is seen as perhaps obscuring the findings rather than assisting in their interpretation.

The opposite viewpoint, in many ways, is taken by the business community. For some, an overall figure of merit is a natural outcome of a human factors audit. With such a figure in hand, the relative needs of different divisions, plants, or departments can be assessed in terms of ergonomic and engineering effort required. Thus, resources can be distributed rationally from a management level. This view is heard by those in the manufacturing and service industries who after an audit ask "How did we do?" and expect a very brief answer. The proliferation of the spreadsheet, with its ability to sum and average rows and columns of data, has encouraged people to do just that with audit results. Repeated audits fit naturally into this view, as they can become the basis for monthly, quarterly, or annual graphs of ergonomic performance.

Neither view alone is entirely defensible. Of course, summing lighting and seating needs produces a result that is logically indefensible and that does not help diagnosis. But equally, decisions must be made concerning optimum use of limited resources. The human factors auditor, having chosen an unbiased sampling scheme and collected data on (presumably) the correct issues, is perhaps in an excellent position to assist in such management decisions. But so, too, are other stakeholders, primarily the workforce.

Audits are not, however, the only use of some of the data collection tools. For example, the Keyserling et al. (1993) upper extremity checklist was developed specifically as a screening tool. Its objective was to find which jobs/workplaces are in need of detailed ergonomic study. In such cases, summing across issues for a total score has an operational meaning (i.e., that a particular workplace needs ergonomic help).

Where interpretation is made at a deeper level than just a single number, a variety of presentation devices have been used. These must show scores (percent of workplaces, distribution of sound pressure levels, etc.) separately but so as to highlight broader patterns. Much is now known about separate versus integrated displays and emergent features (e.g., Wickens, 1992, pp. 121–122), but the traditional profiles and spider's web charts are still the most usual presentation forms. Thus, Wagner (1989) shows the AVISEM profile for a steel industry job before and after automation. The nine different issues ("rating factors") are connected by lines to show emergent shapes for the old and the new jobs. Landau and Rohmert's (1981) original book on AET shows many other examples of profiles. Klimer et al. (1989) present a spider web diagram to show how three work structures influenced 10 issues from the AET analysis. Mattila and Kivi (1989) present their data on the job load and hazard analysis system applied to the building industry in the form of a table. For six occupations, the rating on five different loads/hazards is presented as symbols of different sizes within the cells of the table.

There is little that is novel in the presentation of audit results: Practitioners tend to use the standard tabular or graphical tools. But audit results are inherently multidimensional, so that some thought is needed if the reader is to be helped toward an informed comprehension of the audit's outcome.

5.4 Audit Systems in Practice

Almost any of the audit programs and checklists referred to in previous sections give examples of their use in practice. Only two examples will be given here, as others are readily accessible. These examples were chosen as they represent quite different approaches to auditing.

5.4.1 Auditing a Decentralized Business

From 1992 to 1996, a major U.S.-based apparel manufacturer had run an ergonomics program aimed primarily at the reduction of workforce injuries in backs and upper extremities. As detailed in Drury et al. (1999), the company during that time comprised nine divisions and employed about 45,000 workers. Of particular interest was the fact that the divisions enjoyed great autonomy, with only a small corporate headquarters with a single executive responsible for all risk management activities. The company had grown through mergers and acquisitions, meaning that different divisions had different degrees of vertical integration. Hence, core functions such as sewing, pressing, and distribution were common to most divisions, while some also included weaving, dyeing, and embroidery. In addition, the products and fabrics presented quite different ergonomic challenges, from delicate undergarments, through heavy jeans, to knitted garments and even luggage.

The ergonomics program was similarly diverse. It started with a corporate launch by the highest level executives, then was rolled out to the divisions and to individual plants. The pace of change was widely variable. All divisions were given a standard set of workplace analysis and modification tools (based on Drury and Wick, 1984) but were encouraged to develop their own solutions to problems in a way appropriate to their specific needs.

Evaluation took place continuously, with regular meetings between representatives of plants and divisions to present results of before-and-after workplace studies. However, there was a need for a broader audit of the entire corporation aimed at understanding how much had been achieved for the multimillion-dollar investment, where the program was strong or weak, and what program needs were emerging for the future. During 1995, a team of auditors visited all nine divisions and a total of 12 plants spread across eight divisions. This was three years after the initial corporate launch and about two years after the start of shop-floor implementation.

A three-part audit methodology was used. First, a workplace survey was developed based on elements of the program itself, supplemented by direct comparisons to ergonomics standards and good practices. Table 10 shows this 50-item survey form, with data added for the percentage of "yes" answers where the responses were not measures or scale values. The workplace survey was given at a total of 157 workplaces across the 12 plants. Second, a user survey (Table 11) was used in an interview format with 66 consumers of ergonomics,

Table 10 Ergonomics Audit: Workplace Survey with Overall Data

Number			Division	Plant	Job Type
	Yes	No		Factor	
1. Postural aspects					
W1	68%			Frequent extreme motions of back, neck, shoulders, wrists	
W2	66%			Elbows raised or unsupported more than 50% of time	
W3	22%			Upper limbs contact nonrounded edges	
W4	73%			Gripping with fingers	
W5	36%			Knee/foot controls	
1.1. Seated					
W6	12%			Leg clearance restricted	
W7	21%			Feet unsupported/legs slope down	
W8	17%			Chair/table restricts thighs	
W9	22%			Back unsupported	
W10	37%			Chair height not adjustable easily	
1.2. Standing					
W11	3%			Control requires weight on one foot more than 50% time	
W12	37%			Standing surface hard	
W13	92%			Work surface height not adjustable easily	
1.3. Hand tools					
W14	77%			Tools require hand/wrist bending	
W15	9%			Tools vibrate	
W16	63%			Restricted to one handed use	
W17	39%			Tool handle ends in palm	
W18	20%			Tool handle has nonrounded edges	
W19	56%			Tool uses only two or three fingers	
W20	9%			Requires continuous or high force	
W21	41%			Tool held continuously in one hand	
2. Vibration					
W22	14%			Vibration reaches body from any source	
3. Manual materials handling					
W23	40%			More than five moves per minute	
W24	36%			Loads unbalanced	
W25	14%			Lift above head	
W26	28%			Lift off floor	
W27	83%			Reach with arms	
W28	78%			Twisting	
W29	60%			Bending trunk	
W30	3%			Floor wet or slippery	
W31	0%			Floor in poor condition	
W32	17%			Area obstructs task	
W33	4%			Protective clothing unavailable	
W34	2%			Handles used	
4. Visual aspects					
W35				Task nature: 1, rough; 2, moderate; 3, fine; 4, very fine	
W36				Glare/reflection: 0, none; 1, noticeable; 2, severe	
W37				Color contrast: 0, none; 1, noticeable; 2, severe	
W38				Luminance contrast: 0, none; 1, noticeable; 2, severe	
W39				Task illuminance (foot candles)	
W40	69%			Luminance: task > midfield > outerfield = yes	
5. Thermal aspects					
W41				Dry-bulb temperature (°F)	
W42				Relative humidity (%)	

(*continues*)

Table 10 (*Continued*)

Number	Yes	No	Division	Plant	Job Type / Factor
W43					Airspeed: 1, just perceptible; 2, noticeable; 3, severe
W44					Metabolic cost
W45					Clothing (clo value)
6. Auditory aspects					
W46					Maximum sound pressure level (dBA)
W47					Noise sources: 1, m/c; 2, other m/c; 3, general; 4, other
7. General factors					
W48					Primary cycle time (seconds)
W49	62%				Seen ergonomics video
W50	38%				Any ergonomics changes to workplace or methods

Table 11 Ergonomics Audit: User Survey

Number	Division	Plant	Job Type

U1. What is ergonomics?
U2. Who do you call to do ergonomics?
U3. When did you last ask them to do ergonomics?
U4. Describe what they did.
U5. Who else should we talk to about ergonomics?
U6. General comments on ergonomics.

Table 12 Ergonomics Audit: Provider Survey

Number	Division	Plant	Job Type

P1. What do you do?
P2. How do you get contacted to do ergonomics?
P3. When were you last asked to do ergonomics?
P4. Describe what you did.
P5. How long have you been doing ergonomics?
P6. How were you trained in ergonomics?
P7. What percent of your time is spent on ergonomics?
P8. Where do you go for more detailed ergonomics help?
P9. What ergonomics implementation problems have you had?
P10. How well are you regarded by management?
P11. How well are you regarded by the workforce?
P12. General comments on ergonomics.

typically plant managers, production managers, human resource managers, or their equivalent at the division level, usually vice presidents. Finally, a total of 27 providers of ergonomics services were given a similar provider survey (Table 12) interview. Providers were mainly engineers, with three human resources specialists and one line supervisor. From these three audit methods the corporation wished to provide a time snapshot of how effectively the current ergonomics program was meeting their needs for reduction of injury costs. While the workplace survey measured how well ergonomics was being implemented at the workplace, the user and provider surveys provided data on the roles of the decision makers beyond the workplace.

Detailed audit results are provided in Drury et al. (1999), so only examples and overall conclusions are covered in this chapter. Workplaces showed some evidence of good ergonomic practice, with generally satisfactory thermal, visual, and auditory environments. There were some significant differences ($p < 0.05$) between workplace types rather than between divisions or plants; for example, better lighting (>700 lux) was associated with inspection and sewing. Also, higher thermal load was associated with laundries and machine load/unload. Overall, 83% of workplaces met the ASHRAE (1990) summer comfort zone criteria. As shown in Table 13, the main ergonomics problem areas were in poor posture and manual materials handling. Where operators were seated (only 33% of all workplaces), seating was relatively good. In fact, many of the workforce had been supplied with well-designed chairs as part of the ergonomics program.

To obtain a broad perspective, the three general factors at the end of Table 10 were analyzed. Apart from cycle time (W48), the questions related to workers having seen the corporate ergonomics video (W49) and having experienced a workplace or methods change (W50). Both should have received a "yes" response if the ergonomics program were reaching the entire workforce. In fact, both showed highly significant differences between plants, $X_8^2 = 92.0$, $p < 0.001$ and $X_8^2 = 22.2$, $p < 0.02$, respectively). Some of these differences were due to two divisions lagging in ergonomics implementation, but even beyond this there were large between-plant differences. Overall, 62% of the workforce had seen the ergonomics video, a reasonable value but one with wide variance between plants and divisions. Also, 38% of workplaces had experienced some change, usually ergonomics related,

Table 13 Responses to Ergonomics User

Question and Concern	Corporate Mgt.	Corporate Staff	Plant Mgt.	Plant Staff
1. What is ergonomics?				
1.1. Fitting job to operator	1	6	10	5
1.2. Fitting operator to job	0	6	0	0
2. Who do you call on to get ergonomics work done?				
2.1. Plant ergonomics people	0	3	3	2
2.2. Division ergonomics people	0	4	5	2
2.3. Personnel department	3	0	0	0
2.4. Engineering department	1	8	6	11
2.5. We do it ourselves	0	2	1	0
2.6. College interns	0	0	4	2
2.7. Vendors	0	0	0	1
2.8. Everyone	0	1	0	0
2.9. Operators	0	1	0	0
2.10. University faculty	0	0	1	0
2.11. Safety	0	1	0	0
3. When did you last ask them for help?				
3.1. Never	0	4	2	0
3.2. Sometimes/infrequently	2	0	1	0
3.3. 1 year or more ago	0	1	4	0
3.4. 1 month or so ago	0	0	2	0
3.5. Less than 1 month ago	1	0	3	4
5. Who else should we talk to about ergonomics?				
5.1. Engineers	0	0	3	2
5.2. Operators	1	1	2	0
5.3. Everyone	0	0	2	0
6. General ergonomics comments				
6.1. Ergonomics concerns				
6.1.1. Workplace design for safety/ease/stress/fatigue	2	5	13	5
6.1.2. Workplace design for cost savings/productivity	1	0	2	1
6.1.3. Workplace design for worker satisfaction	1	1	0	1
6.1.4. Environment design	2	1	3	0
6.1.5. The problem of finishing early	0	0	1	1
6.1.6. The seniority/bumping problem	0	3	1	0
6.2. Ergonomics program concerns				
6.2.1. Level of reporting of ergonomics	0	1	7	0
6.2.2. Communication/who does ergonomics	7	1	4	0
6.2.3. Stability/staffing of ergonomics	0	0	10	4
6.2.4. General evaluation of ergonomics				
Positive	1	3	3	4
Negative	4	10	10	3
6.2.5. Lack of financial support for ergonomics	0	0	1	0
6.2.6. Lack of priority for ergonomics	2	2	1	4
6.2.7. Lack of awareness of ergonomics	2	1	6	1

a respectable figure after only two to three years of the program.

From the user and provider surveys, an enhanced picture emerged. Again, there was variability between divisions and plants, but 94% of the users defined ergonomics as fitting the job to the operator rather than training or medical management of injuries. Most users had requested an ergonomic intervention within the past two months, but other "users" had never in fact used ergonomics.

The solutions employed ranged widely, with a predominance of job aids such as chairs or standing pads. Other frequent categories were policy changes (e.g., rest breaks, rotation, box weight reduction) and workplace adjustment to the individual operator. There were few uses of personal aids (e.g., splints) or referrals to physicians as ergonomic solutions. Changes to the workplace clearly predominated over changes to the individual, although a strong medical management program was in place when required. When questioned about ergonomics results, all mentioned safety (or workplace comfort or ease of use), but some also mentioned others. Cost or productivity benefits were the next most common response, with a few additional ones relating to employee relations, absence/turnover, or job satisfaction. Significantly, only one respondent mentioned quality.

The major user concern at the plant level was time devoted to ergonomics by providers. At the corporate level, the need was seen for more rapid job analysis methods and corporate policies (e.g., on back belts or "good" chairs). Overall, 94% of users made positive comments about the ergonomics program.

Ergonomics providers were almost always trained in the corporate or division training seminars, usually near the start of the program. Providers' chief concern was for the amount of time and resources they could spend on ergonomics activities. Typically, ergonomics was only one job responsibility among many. Hence, broad programs, such as new chairs or back belts, were supported enthusiastically, as they gave the maximum perceived impact for the time devoted. Other solutions presented included job aids, workplace redesign (e.g., moving from seated to standing jobs for long-seam sewing), automation, rest breaks, job rotation, packaging changes, and medical management. Specific needs were seen in the area of corporate or supplier help in obtaining standard equipment solutions and of more division-specific training. As with users, the practitioners enjoyed their ergonomics activity and thought it worthwhile.

Recommendations arising from this audit were that the program was reasonably effective at present but had some long-term needs. The corporation sees itself as an industry leader and wants to move beyond a relatively superficial level of ergonomics application. To do this will require more time resources for job analysis and change implementation. Corporate help could also be provided in developing more rapid analysis methods, standardized video-based training programs, and more standardized solutions to recurring ergonomics problems. Many of these changes have since been implemented.

On another level, the audit was a useful reminder to the company of the fact that it had incurred most of the up-front costs of a corporate ergonomics program and was now beginning to reap the benefits. Indeed, by 1996, corporate injury costs and rates had decreased by about 20% per year after peaking in 1993. Clearly, the ergonomics program was not the only intervention during this period, but it was seen by management as the major contributor to improvement. Even on the narrow basis of cost savings, the ergonomics program was a success for the corporation.

5.4.2 Error Reduction at a Colliery

In a two-year project reported by Fox (1992) and Simpson (1994), the human error audit described in Section 5.3.2 was applied to two colliery haulage systems. The results of the first study are presented here. In both systems, data collection focused on potential errors and the performance-shaping factors (PSFs) that can influence these errors. Data were collected by "observation, discussion and measurement within the framework of the broader man–machine systems and checklist of PSFs," taking some 30–40 shifts at each site. The entire haulage system from surface operations to delivery at the coal face was covered.

The first study found 40 active failures (i.e., direct error precursors) and 9 latent failures (i.e., dormant states predisposing the system to later errors). Four broad classes of active failures were (1) errors associated with locomotive maintenance (7 errors) (e.g., fitting incorrect thermal cutoffs), (2) errors associated with locomotive operation (10 errors) (e.g., locomotives not returned to the service bay for a 24-h check), (3) errors associated with loads and load security (7 errors), (e.g., failure to use spacer wagons between overhanging loads), and (4) errors associated with the design/operation of the haulage route (10 errors), (e.g., continued use despite potentially unsafe track) plus a small miscellaneous category.

The latent failures were (Fox, 1992) (1) quality assurance in supplying companies, (2) supply-ordering procedures within the colliery, (3) locomotive design, (4) surface "makeup" of supplies, (5) lack of equipment at specific points, (6) training, (7) attitudes to safety, and (8) the safety inspection/reporting/action procedures. As an example of item 3, locomotive design, the control positions were not consistent across the locomotive fleet, despite all originating from the same manufacturer. Using the slip/mistake/violation categorization, each potential error could be classified so that the preferred source of action (intervention) could be specified.

This audit led to the formation of two teams, one to tackle locomotive design issues and the other for safety reporting and action. As a result of team activities, many ergonomic actions were implemented. These included management actions to ensure a uniform wagon fleet, autonomous inspection/repair teams for tracks, and multifunctional teams for safety initiatives.

The outcome was that the accident rate dropped from 35.40 per 100,000 person-shifts to 8.03 in one year. This brought the colliery from worst in the regional group of 15 collieries to best in the group and indeed in the United Kingdom. In addition, personnel indicators, such as industrial relations climate and absence rates, improved.

6 CONCLUSIONS

In this chapter we have arrived at human factors audits through a context of inspection and checklist design. It should be obvious by now that checklists are a subset of audits, which are in turn a subset of inspection. Within the context of inspection, we have seen that all inspections follow a short logical sequence of functions and that each function has considerable scope for model-based and empirical design to improve the human factors and system performance. Nonmanufacturing applications have been emphasized, with the focus on processes and broader systems rather than on repetitively produced products. Audits have been shown to be functionally similar to inspections.

Inspecting, checking, and auditing are interesting, as they all have human factors design aspects but can all be applied to both the processes being audited and the auditing process itself. Whether inspecting nonmanufacturing items or checking items on a checklist or performing an audit, there is prescriptive advice on how to develop or choose a system that accords with human factors good practices.

Disclaimer: The findings and conclusions in this chapter are those of the authors and do not necessarily represent the views of the National Institute for Occupational Safety and Health.

REFERENCES

Alexander, D. C., and Pulat, B. M. (1985), *Industrial Ergonomics: A Practitioner's Guide*, Industrial Engineering and Management Press, Atlanta, GA.

American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE) (1989), "Physiological Principles, Comfort and Health," in *Fundamentals Handbook*, ASHRAE, Atlanta, GA, Chapter 8.

American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE) (1990), *Handbook of Fundamentals*, ASHRAE, Atlanta, GA.

Arani, T., Karwan, M. H., and Drury, C. G. (1984), "A Variable-Memory Model of Visual Search," *Human Factors*, Vol. 26, No. 6, pp. 631–639.

AVISEM (1977), *Techniques d'amélioration des conditions de travail dans l'industrie*, Editions Hommes et Techniques, Suresnes, France.

Ayoub, M. M., and Mital, A. (1989), *Manual Materials Handling*, Taylor & Francis, London.

Bainbridge, L., and Beishon, R. J. (1964), "The Place of Checklists in Ergonomic Job Analysis," in *Proceedings of the 2nd I.E.A. Congress*, Dortmund, Ergonomics Congress Proceedings Supplement.

Baveja, A., Drury, C. G., Karwan, M., and Malone, D. M. (1996), "Derivation and Test of an Optimum Overlapping-Lobe Model of Visual Search," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 28, pp.161–168.

Berchem-Simon, O., Ed. (1993), *Ergonomics Action in the Steel Industry*, EUR 14832 EN, Commission of the European Communities, Luxembourg.

Blackwell, H. R. (1946), "Contrast Thresholds of the Human Eye," *Journal of the Optical Society of America*, Vol. 36, pp. 624–643.

Bolijn, A. J. (1993), *Research into the Employability of Women in Production and Maintenance Jobs in Steelworks*, Commission of the European Communities, Luxembourg.

British Standards Institution (BSI) (1980), "Educational furniture," *Specification for Functional Dimensions, Identification and Finish of Chairs and Tables for Educational Institutions*, BS5873-1: 1980, BSI, London.

Burger, G. C. E., and de Jong, J. R. (1964), "Evaluation of Work and Working Environment in Ergonomic Terms," *Aspects of Ergonomic Job Analysis*, Vol. 7, pp. 185–201.

Cakir, A., Hart, D. M., and Stewart, T. F. M. (1980), *Visual Display Terminals*, Wiley, New York.

Carson, A. B., and Carlson, A. E. (1977), *Secretarial Accounting*, 10th ed., South-Western, Cincinnati, OH.

Chan, A. H. S., and Chiu, C. H. Y., (2010), "Visual Lobe Shape Characteristics of Experienced Inspectors and Inexperienced Subjects," *Human Factors and Ergonomics in Manufacturing and Service Industries*, Vol. 20.5, pp. 367–377.

Chen, Y., Gale, A., and Scott, H. (2008), "Anytime, Anywhere Mammographic Interpretation Training," in *Contemporary Ergonomics 2008*, P. D. Bust, Ed., Taylor & Francis, London, pp. 375–380

Chervak, S., and Drury, C. G. (1995), "Simplified English Validation," in *Human Factors in Aviation Maintenance: Phase 6 Progress Report*, DOT/FAA/AM-95/xx, Federal Aviation Administration, Office of Aviation Medicine, National Technical Information Service, Springfield, VA.

Chervak, S., Drury, C. G., and Ouellette, J. L. (1996), "Simplified English for Aircraft Workcards," in *Proceedings of the Human Factors and Ergonomics Society 39th Annual Meeting*, pp. 303–307.

Chi, C.-F., and Drury, C. G. (1995), "A Test of Economic Models of Stopping Policy in Visual Search," *IIE Transactions*, Vol. 27, pp. 392–393.

Croft, J. (2001). Research perfect new ways to monitor pilot performance. *Aviation Week & Space Technology*, 155, 76.

Degani, A., and Wiener, E. L. (1990), *Human Factors of Flight-Deck Checklists: The Normal Checklist*, Contractor Report 177548, NASA Ames Research Center, Moffett Field, CA.

Dempsey, P. G., McGorry, R. W., and Maynard, W. S. (2005), "A Survey of Tools and Methods Used by Certified Professional Ergonomists," *Applied Ergonomics*, Vol. 36, No. 4, pp. 489–503.

Devor, R. E., Change, T.-H., and Sutherland, J. W. (1992), *Statistical Quality Design and Control*, Prentice-Hall, Englewood Cliffs, NJ.

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

Drury, C. G. (1978), "Integration of Human Factors Models into Statistical Quality Control," *Human Factors*, Vol. 20, No. 5, pp. 561–572.

Drury, C. G. (1990a), "The Ergonomics Audit," in *Contemporary Ergonomics*, E. J. Lovesey, Ed., Taylor & Francis, London, pp. 400–405.

Drury, C. G. (1990b), "Computerized Data Collection in Ergonomics," in *Evaluation of Human Work*, J. R. Wilson and E. N. Corlett, Eds., Taylor & Francis, London, pp. 200–214.

Drury, C. G. (1992a), "Inspection Performance," in *Handbook of Industrial Engineering*, G. Salvendy, Ed., Wiley, New York, pp. 2282–2314.

Drury, C. G. (1992b), "Design for Inspectability," in *Design for Manufacturability: A Systems Approach to Concurrent Engineering and Ergonomics*, M. H. Helander and M. Nagamachi, Eds., Taylor & Francis, London, pp. 204–216.

Drury, C. G. (1997), "Ergonomics Audits: Why and How," in *Proceedings of the 13th Triennial Congress of the International Ergonomics Association'97*, Vol. 2, Tampere, Finland, pp. 284–286.

Drury, C. G. (2002), "A Unified Model of Security Inspection," in *Proceedings of the Aviation Security Technical Symposium*, Federal Aviation Administration, Atlantic City, NJ, pp. 27–30.

Drury, C. G. (2003, April–May), "Service Quality and Human Factors," *AI and Society*, Vol. 17, No. 2, pp. 78–96.

Drury, C. G. (2009), "Inspection: From Goods through Passengers to Systems," in P. D. Bust (Ed) *Contemporary Ergonomics 2009*, Taylor and Francis, London, pp. 3–15.

Drury, C. G., and Addison, J. L. (1973), "An Industrial Study of the Effects of Feedback and Fault Density on Inspection Performance," *Ergonomics*, Vol. 16, pp. 159–169.

Drury, C. G., and Sinclair, M. A. (1983), "Human and Machine Performance in an Inspection Task," *Human Factors*, Vol. 25, No. 4, pp. 391–400.

Drury, C. G., and Wick, J. (1984), "Ergonomic Applications in the Shoe Industry," in *Proceedings of the International Conference on Occupational Ergonomics*, Vol. 1, pp. 489–483.

Drury, C. G., Kleiner, B. M., and Zahorjan, J. (1989), "How Can Manufacturing Human Factors Help Save a Company: Intervention at High and Low Levels," in *Proceedings of the Human Factors Society 33rd Annual Meeting*, Denver, CO, pp. 687–689.

Drury, C. G., Prabhu, P., and Gramopadhye, A. (1990), "Task Analysis of Aircraft Inspection Activities: Methods and Findings," in *Proceedings of the Human Factors Society 34th Annual Meeting*, Santa Monica, CA, pp. 1181–1185.

Drury, C. G., Spencer, F. W., and Schurman, D. (1997), "Measuring Human Detection Performance in Aircraft Visual Inspection," in *Proceedings of the 41st Annual Human Factors and Ergonomics Society Meeting*, Albuquerque, NM.

Drury, C. G., Broderick, R. L., Weidman, C. H., and Reynolds Mozrall, J. L. (1999), "A Corporate-Wide Ergonomics Programme: Implementation and Evaluation," *Ergonomics*, Vol. 42, No. 1, pp. 208–228.

Drury, C. G., Patel, S. C., and Prabhu, P. V. (2000), "Relative Advantage of Portable Computer-Based Workcards for Aircraft Inspection," *International Journal of Industrial Ergonomics*, Vol. 26, pp. 163–176.

Drury, C. G., Ghylin, K. M., and Schwaninger, A. (2007), "Large-Scale Validation of a Security Inspection Model," in P. D. Bust (Ed) *Contemporary Ergonomics 2007*, Taylor & Francis, London

Easterby, R. S. (1967), "Ergonomics Checklists: An Appraisal," *Ergonomics*, Vol. 10, No. 5, pp. 548–556.

Engel, F. L. (1971), "Visual Conspicuity, Directed Attention, and Retinal Locus," *Visual Research*, Vol. 11, pp. 563–576.

Eriksen, C. W. (1990), "Attentional Search of the Visual Field," in *Visual Search*, D. Brogan, Ed., Taylor & Francis, London, pp. 3–19.

Evans, J. R., and Lindsay, W. (1993), *The Management and Control of Quality*, West Publishing Company, Minneapolis/St. Paul, MN.

Fanger, P. O. (1970), *Thermal Comfort, Analyses and Applications in Environmental Engineering*, Danish Technical Press, Copenhagen, Denmark.

Federal Aviation Administration (FAA) (2006), "Line Operations Safety Audits," Advisory Circular AC 120-90, FAA, Washington, DC.

Fox, J. G. (1992), "The Ergonomics Audit as an Everyday Factor in Safe and Efficient Working," *Progress in Coal, Steel and Related Social Research*, Vol. 11, pp. 10–14.

Gale, A. G., Mugglestone, M. D., Purdy, K. J., and McClumpha, A. (2000), "Is Airport Baggage Inspection Just Another Medical Image?," *Proceedings of the SPIE*, Vol. 3981, pp. 184–192.

Ghylin, K.M., Drury, C.G., Batta, R. & Lin, L. (2007). Temporal Effects in a Security Inspection Task: Breakdown of Performance Components. *Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting*, Oct. 1-5, 2007, 93–97, Baltimore, MD

Givoni, B., and Goldman, R. F. (1972), "Predicting Rectal Temperature Response to Work, Environment, and Clothing," *Journal of Applied Physiology*, Vol. 32, No. 6, pp. 812–822.

Goldberg, J. H., and Gibson, D. C. (1986), "The Effects of Training Method and Type of Checklist upon Visual Inspection Accuracy," in *Proceedings of the Annual International Industrial Ergonomics and Safety Conference*, Louisville, KY, North-Holland, Amsterdam.

Groth, K. M. (1989), "The Modular Work Analysis System (MAS)," in *Recent Developments in Job Analysis, Proceedings of the International Symposium on Job Analysis*, Taylor & Francis, New York, pp. 253–261.

Guélaud, F., Beauchesne, M.-N., Gautrat, J., and Roustang., G. (1975), "Pour une analyse des conditions de travail ouvrier dans l'entreprise," in *Recherche du Laboratoire d'Economie et de Sociologie du Travail C.M.R.S. Aix-en-Provence*, 3rd ed., Librairie Armand Colin, Paris.

Hacker, W., Iwanowa, A., and Richter, P. (1983), "Tätigkeitsbewertungssystem," paper presented at the Psychodiagnostisches Zentrum, Berlin (OST).

Hasslequist, R. J. (1981), "Increasing Manufacturing Productivity Using Human Factors Principles," in *Proceedings of the Human Factors Society 25th Annual Meeting*, Santa Monica, CA, pp. 204–206.

Helmreich, R. L., Wilhelm, J. A., Klinect, J. R., and Merritt, A. C. (2001), "Culture, Error and Crew Resource Management," in E. Salas, C. A. Bowers, and E. Edens, Eds., *Applying Resource Management in Organizations: A Guide for Professionals*, Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 305–331.

Hong, S.-K., and Drury, C. G. (2002), "Sensitivity and Validity of Visual Search Models for Multiple Targets," *Theoretical Issues in Ergonomic Science*, Vol. 3, No. 1, pp. 1–26.

Hou, T.-S., Lin, L., and Drury, C. G. (1993), "An Empirical Study of Hybrid Inspection Systems and Allocation of Inspection Function," *International Journal of Human Factors in Manufacturing*, Vol. 3, pp. 351–367.

Hsiao, Y-L, Joseph, C., Joshi, S., Lapp, J., Pennathur, P. R., and Drury, C. G. (2008), "Posture and Performance: Sitting vs. Standing for Security Screening," *Ergonomics*, Vol. 51, pp. 290–307.

Illuminating Engineering Society (IES) (1993), *Lighting Handbook, Reference and Application*, 8th ed., IES of North America, New York.

International Civil Aviation Organization (ICAO) (1989), *Human Factors Digest No. 1 Fundamental Human Factors Concepts*, Circular 216-AN/131, ICAO, Montreal, Canada.

International Labour Office (ILO) (1999), *Ergonomic Checkpoints*. ILO, Geneva.

Jamieson, G. H. (1966), "Inspection in the Telecommunications Industry: A Field Study of Age and Other Performance Variables," *Ergonomics*, Vol. 9, pp. 297–303.

Jiang, X., Melloy, B. J., and Gramopadhye, A. (2003), "Evaluation of Best System Performance: Human, Automated and Hybrid Inspection Systems," *Human Factors and Ergonomics in Manufacturing*, Vol. 13, No. 2, pp. 137–152.

Jones, J. C. (1981), *Design Methods: Seeds of Human Factors*, Wiley, New York.

Jones, D. M., and Broadbent, D. E. (1987), "Noise," in *Handbook of Human Factors Engineering*, G. Salvendy, Ed., Wiley, New York.

Karwowski, W., (2005), *Handbook of Standards and Guidelines in Ergonomics and Human Factors (Human Factors/Ergonomics)*, Lawrence Erlbaum Associates, Mahwah, NJ.

Kerlinger, F. N. (1964), *Foundations of Behavioral Research*, Holt, Rinehart and Winston, New York.

Keyserling, W. M., Brouwer, M., and Silverstein, B. A. (1992), "A Checklist for Evaluation of Ergonomic Risk Factors Resulting from Awkward Postures of the Legs, Trunk and Neck," *International Journal of Industrial Ergonomics*, Vol. 9, No. 4, pp. 283–301.

Keyserling, W. M., Stetson, D. S., Silverstein, B. A., and Brouwer, M. L. (1993), "A Checklist for Evaluating Ergonomic Risk Factors Associated with Upper Extremity Cumulative Trauma Disorders," *Ergonomics*, Vol. 36, No. 7, pp. 807–831.

Kirwan, B. (1989), "A Human Factors and Human Reliability Programme for the Design of a Large UK Nuclear Chemical Plant," in *Proceedings of the Human Factors Society 33rd Annual Meeting*, Denver, CO, pp. 1009–1013.

Klimer, F., Kylian, H., Schmidt, K.-H., and Rutenfranz, J. (1989), "Work Analysis and Load Components in an Automobile Plant after the Implementation of New Technologies," in *Recent Developments in Job Analysis*, K. Landau and W. Rohmert, Eds., Taylor & Francis, New York, pp. 331–340.

Klinect, J. R. (2005), Line "Operations Safety Audit: A Cockpit Observation Methodology for Monitoring Commercial Airline Safety Performance," Ph.D. Dissertation, University of Texas at Austin, 2005

Klinect, J. R., Murray, P., Merritt, A., and Helmreich, R. (2003), "Line Operations Safety Audit (LOSA): Definition and Operating Characteristics," in *Proceedings of the 12th International Symposium on Aviation Psychology*, The Ohio State University, Dayton, OH, pp. 663–668.

Kogi, K. (1994), "Introduction to WISE (Work Improvement in Small Enterprises) Methodology and Workplace Improvements Achieved by the Methodology in Asis," in S. McFadden, L. Innes, and M. Hill, Eds., *Proceedings of the 12th Triennial Congress of the International Ergonomics Association*, Vol. 5, Human Factors Association of Canada, Toronto, Ontario, Canada, pp. 141–143.

Kogi, K., and Kuorinka, I., Eds. (1995), *Ergonomic Checkpoints*, ILO Publications, Geneva, Switzerland.

Koli, S., Drury, C. G., Cuneo, J., and Lofgren, J. (1993), "Ergonomic Audit for Visual Inspection of Aircraft," in *Human Factors in Aviation Maintenance: Phase Four, Progress Report*, DOT/FAA/AM-93/xx, National Technical Information Service, Springfield, VA.

Koli, S. T. (1994), "Ergonomic Audit for Non-Repetitive Task," M.S. Thesis, State University of New York at Buffalo, Buffalo, NY.

Lamar, E. S. (1960), "Operational Background and Physical Considerations Relative to Visual Search Problems," in *Visual Search Techniques*, A. Morris and P. E. Horne, Eds., National Research Council, Washington, DC.

Landau, K., and Rohmert, W. (1981), *Fallbeispiele zur Arbeitsanalyse*, Hans Huber, Bern, Switzerland.

Landau, K., and Rohmert, W., Eds. (1989), *Recent Developments in Job Analysis: Proceedings of the International Symposium on Job Analysis*, University of Hohenheim, March 14–15, Taylor & Francis, New York.

Larock, B. (2000), "The Effects of Spatial and Functional Job-Aid Checklists on a Frequently Performed Aircraft Maintenance Task," Master's Thesis, State University of New York at Buffalo, Buffalo, NY.

Larock, B., and Drury, C. G. (2003), "Repetitive Inspection with Checklists: Design and Performance," in *Proceedings of the Ergonomics Society Annual Conference*, Edinburgh, Scotland, April 15–17.

Leitner, K., and Greiner, B. (1989), "Assessment of Job Stress: The RHIA Instrument," in *Recent Developments in Job Analysis, Proceedings of the International Symposium on Job Analysis*, K. Landau and W. Rohmert, Eds., University of Hohenheim, March 14–15, 1989, Taylor & Francis, New York.

Lloyd, C. J., Boyce, P., Ferzacca, N., Eklund, N., and He, Y. (2000), "Paint Inspection Lighting: Optimization of Lamp Width and Spacing," *Journal of the Illuminating Engineering Society,* Vol. 28, pp. 99–102

Ma, J., Pedigo, M., Blackwell, L., Hackworth, C., Holcomb, K., and Gildea, K. and Hiles, J. J., (2011), "The Line Operations Safety Audit Program:Transitioning rom Flight Operations to Maintenance and Ramp Operations," DOT/FAA/AM-11/15, Office of Aerospace Medicine, Washington, DC 20591

Malde, B. (1992), "What Price Usability Audits? The Introduction of Electronic Mail into a User Organization," *Behaviour and Information Technology,* Vol. 11, No. 6, pp. 345–353.

Malone, T. B., Baker, C. C., and Permenter, K. E. (1988), "Human Engineering in the Naval Sea Systems Command," in *Proceedings of the Human Factors Society 32nd Annual Meeting*, Anaheim, CA, Vol. 2, pp. 1104–1107.

Mattila, M., and Kivi, P. (1989), "Job Load and Hazard Analysis: A Method for Hazard Screening and Evaluation," in *Recent Developments in Job Analysis: Proceedings of the International Symposium on Job Analysis*, K. Landau and W. Rohmert, Eds., University of Hohenheim, March 14–15, Taylor & Francis, New York, pp. 179–186.

McClelland, I. (1990), "Product Assessment and User Trials," in *Evaluation of Human Work*, J. R. Wilson and E. N. Corlett, Eds., Taylor & Francis New York, pp. 218–247.

McCormick, E. J. (1979), *Job Analysis: Methods and Applications*, AMACOM, New York.

McCormick, W. T., Mecham, R. C., and Jeanneret, P. R. (1969), *The Development and Background of the Position Analysis Questionnaire*, Occupational Research Center, Purdue University, Lafayette, IN.

McKenzie, R. M. (1958), "On the Accuracy of Inspectors," *Ergonomics*, Vol. 1, No. 3, pp. 258–272.

Miller, R. B. (1967), "Task Taxonomy: Science or Art," in *The Human Operator in Complex Systems*, W. T. Singleton, R. S. Easterby, and D. Whitfield, Eds., Taylor & Francis, London.

Mir, A. H. (1982), "Development of Ergonomic Audit System and Training Scheme," M.S. Thesis, State University of New York at Buffalo, Buffalo, NY.

Morawski, T., Drury, C. G., and Karwan, M. H. (1980), "Predicting Search Performance for Multiple Targets," *Human Factors*, Vol. 22, No. 6, pp. 707–718.

National Institute for Occupational Safety and Health (NIOSH) (1981), *Work Practices Guide for Manual Lifting*, DHEW-NIOSH Publication 81–122, NIOSH, Cincinnati, OH.

Nery, D. (1999), "Audit Tool User Guide: Ergonomic Hazard Management for the Meat Industry in South Australia," WorkCover Corporation, Adelaide, Australia, available: http://www.safework.sa.gov.au/contentPages/docs/meatUserGuide.pdf.

Nodine, C. F., Mello-Thoms, C., Kundel, H. L., and Weinstein, S. P. (2002), "Time Course of Perception and Decision Making during Mammographic Interpretation," *American Journal of Roentgenology*, Vol. 179, pp. 917–923.

Occupational Safety and Health Administration (OSHA) (1990), *Ergonomics Program Management Guidelines for Meatpacking Plants*, OSHA-3121, OSHA, U.S. Department of Labor, Washington, DC.

Ockerman, J. J., and Pritchett, A. R. (1998), "Preliminary Investigation of Wearable Computers for Task Guidance in Aviation Inspection," paper presented at the Second International Symposium on Wearable Computers, Pittsburgh, PA.

Ockerman, J. J., and Pritchett, A. R. (2000), "Reducing Over-Reliance on Task-Guidance Systems," in *Proceedings of the Human Factors and Ergonomics Society 44th Annual Meeting*, San Diego, CA.

Ockerman, J. J., and Pritchett, A. R. (2004), "Improving Performance on Procedural Tasks through Presentation of Locational Procedure Context: An Empirical Evaluation," *Behaviour & IT*, Vol. 23, No. 1, pp. 11–20.

Osburn, H. G. (1987), "Personnel Selection," in G. Salvendy, Ed., *Handbook of Human Factors*, Wiley, New York, pp. 911–933.

Overington, I. (1973), *Vision and Acquisition*, Pentech Press, London.

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

Panter, W. (1993), "Biomechanical Damage Risk in the Handling of Working Materials and Tools—Analysis, Possible Approaches and Model schemes," in *Ergonomics Action in the Steel Industry EUR 14832 EN*, O. Berchem-Simon, Ed., Commission of the European Communities, Luxembourg.

Park, J.-K., Boyer, J., Tessler, J., Casey, J., Schemm, L., Gore, R., Punnett, L., and Promoting Healthy and Safe Employment (PHASE) in Healthcare Project Team (2009), "Inter-Rater Reliability of PATH Observations for Assessment of Ergonomic Risk Factors in Hospital Work," *Ergonomics*, Vol. 52, pp. 820–829.

Parsons, K. C. (1992), "The Thermal Audit: A Fundamental Stage in the Ergonomics Assessment of Thermal Environment," in *Contemporary Ergonomics 1992*, E. J. Lovesey, Ed., Taylor & Francis, London, pp. 85–90.

Patel, S., Drury, C. G., and Prabhu, P. (1993), "Design and Usability Evaluation of Work Control Documentation," in *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting*, Seattle, WA, pp. 1156–1160.

Patel, S., Drury, C. G., and Lofgren, J. (1994), "Design of Workcards for Aircraft Inspection," *Applied Ergonomics*, Vol. 25, No. 5, pp. 283–293.

Pearl, A., and Drury, C. G. (1995), "Improving the Reliability of Maintenance Checklists," in *Human Factors in Aviation Maintenance, Phase Four, Progress Report*, DOT/FAA/AM-93/xx, National Technical Information Service, Springfield, VA.

Portillo Sosa, J. (1993), "Design of a Computer Programme for the Detection and Treatment of Ergonomic Factors at Workplaces in the Steel Industry," in *Ergonomics Action in the Steel Industry EUR 14832 EN*, O. Berchem-Simon, Ed., Commission of the European Communities, Luxembourg, pp. 421–427.

Public, L., Wilson, J., and Lowe, E. (2010), "The Operational Demand Evaluation Checklist (ODEC) of Workload for Rail Signalling," *Applied Ergonomics*, Vol. 41, pp. 393–402.

Pulat, B. M. (1992), *Fundamentals of Industrial Ergonomics*, Prentice-Hall, Englewood Cliffs, NJ.

Rasmussen, J. (1987), "Reasons, Causes and Human Error," in *New Technology and Human Error*, J. Rasmussen, K. Duncan, and J. Leplat, Eds., Wiley, New York, pp. 293–301.

Reason, J. (1990), *Human Error*, Cambridge University Press, New York.

Régie Nationale des Usine Renault (RNUR) (1976), "Les profils de ostes: Méthode d'analyse des conditions de travail," in *Collection Hommes et Savoir*, Masson, Sirtés, Paris.

Robson, J. W., and Wolstenhulme, S. (2010), "Ergonomically Auditing Your Scan Room," *Ultrasound*, Vol. 18, pp. 41–44.

Robson, L. S., Macdonald, S., Van Erd, D. L., Gray, G. C., and Bigelow, P. (2010), "Something Might be Missing from Occupational Health and Safety Audits: Findings from a Content Validity Analysis of Five Audit Instruments," *Journal of Occupational and Environmental Medicine*, Vol. 52, pp. 536–543.

Rohmert, W., and Landau, K. (1983), *A New Technique for Job Analysis*, Taylor & Francis, London.

Rohmert, W., and Landau, K. (1989), "Introduction to Job Analysis," *A New Technique for Job Analysis*, Taylor & Francis, London, Part 1, pp. 7–22.

Seeber, A., Schmidt, K.-H., Kierswelter, E., and Rutenfranz, J. (1989), "On the Application of AET, TBS and VERA to Discriminate between Work Demands at Repetitive Short Cycle Tasks," in *Recent Developments in Job Analysis*, K. Landau and W. Rohmert, Eds., Taylor & Francis, New York, pp. 25–32.

SHARE (1990), *Inspecting the Workplace*, SHARE Information Booklet, Occupational Health and Safety Authority, Canberra, Australia.

Simpson, G. C. (1994), "Ergonomic Aspects in Improvement of Safe and Efficient Work in Shafts," in *Ergonomics Action in Mining*, Eur 14831, Commission of the European Communities, Luxembourg, pp. 245–256.

Sinclair, M. A. (1984), "Ergonomics of Quality Control," workshop document, International Conference on Occupational Ergonomics, Toronto, Ontario, Canada.

Sinclair, M. A. (1990), "Subjective Assessment," in *Evaluation of Human Work*, J. R. Wilson and E. N. Corlett, Eds., Taylor & Francis, London pp. 58–88.

Smith, R. T. (2001), "An Ergonomics Lineside Audit Method in Automotive Manufacturing," in D. C. Alexander and R. A. Rabourn, Eds., *Applied Ergonomics*, Taylor & Francis, London, pp. 3–13.

Thomas, L. F., and Seaborne, A. E. M. (1961), "The Sociotechnical Context of Industrial Inspection," *Occupational Psychology*, Vol. 35, pp. 36–43.

Treisman, A. (1986), "Features and Objects in Visual Processing," *Scientific American*, Vol. 255, pp. 114B–125.

U.S. DoD (1989), Human Engineering Design Criteria for Military Systems, Equipment and Facilities, MIL-STD-1472D, U. S. Department of Defense, Washington, DC.

Village, J., Trask, C., Luong, N., Chow, Y., Johnson, P., Koehoorn, M., and Teschke, K. (2009), "Development and Evaluation of an Observational Back-Exposure Sampling Tool (Back-EST) for Work-Related Back Injury Risk Factors," *Applied Ergonomics*, Vol. 40, pp. 538–544.

Wagner, R. (1989), "Standard Methods Used in French-Speaking Countries for Workplace Analysis," in *Recent Developments in Job Analysis*, K. Landau and W. Rohmert, Eds., Taylor & Francis, New York, pp. 33–42.

Wang, M. J., and Drury, C. G. (1989), "A Method of Evaluating Inspector's Performance Differences and Job Requirements," *Applied Ergonomics*, Vol. 20, No. 3, pp. 181–190.

Waters, T. R., Putz-Anderson, V., Garg, A., and Fine, L. J. (1993), "Revised NIOSH Equation for the Design and Evaluation of Manual Lifting Tasks," *Rapid Communications, Ergonomics, 1993*, Vol. 36, No. 7, pp. 748–776.

Wenner, C., and Drury, C. G. (1996), "Active and Latent Failures in Aircraft Ground Damage Incidents," in *Proceedings of the Human Factors and Ergonomic Society 39th Annual Meeting*, pp. 796–799.

Wenner, C., and Drury, C. G. (2000), "Analyzing Human Error in Aircraft Ground Damage Incidents," *International Journal of Industrial Ergonomics,* Vol. 26, No. 2, pp. 177–199.

Westwater, M. G., and Johnson, G. I. (1995), "Comparing Heuristic, User-Centred and Checklist-Based Evaluation Approaches," in *Contemporary Ergonomics 1995*, S. A. Robertson, Ed., Taylor & Francis, London, pp. 538–543.

Wickens, C. D. (1992), *Engineering Psychology and Human Performance*, 2nd ed., Harper-Collins, New York.

Wilkins, J. R., Bean, T. L, Mitchell, G. L., Crawford, J. M., and Eicher, L. C. (1997), "Development and Application of a Pen-Based Computer Program for Direct Entry of Agricultural Hazard Data," *Applied Occupational and Environmental Hygiene*, Vol. 12, No. 2, pp. 105–110.

Wolfe, J. M. (1994), "Guided Search II: A Revised Model of Visual Search," *Psychonomic Bulletin and Review 1994*, Vol. 1, No. 2, pp. 202–238.

Wright, P., and Barnard, P. (1975), "Just Fill in This Form: A Review for Designers," *Applied Ergonomics*, Vol. 6, pp. 213–220.

Yoshida, H., and Ogawa, K. (1991), "Workplace Assessment Guideline—Checking Your Workplace," in *Advances in Industrial Ergonomics and Safety III*, W. Karwowski and J. W. Yates, Eds., Taylor &w Francis, London, pp. 23–28.