HANDBOOK OF HUMAN FACTORS AND ERGONOMICS

Third Edition

Edited by

Gavriel Salvendy
Purdue University
West Lafayette, Indiana
and
Tsinghua University
Beijing, People's Republic of China

WILEY
JOHN WILEY & SONS, INC.

This book is printed on acid-free paper. ♾

Published by John Wiley & Sons, Inc., Hoboken, New Jersey
Published simultaneously in Canada

Disclaimer

The editor, authors, and the publisher have made every effort to provide accurate and complete information in the Handbook but the Handbook is not intended to serve as a replacement for professional advice. Any use of the information in this Handbook is at the reader's discretion. The editor, authors, and the publisher specifically disclaim any and all liability arising directly or indirectly from the use or application of any information contained in this Handbook. An appropriate professional should be consulted regarding your specific situation.

Library of Congress Cataloging-in-Publication Data:

Handbook of human factors and ergonomics / edited by Gavriel Salvendy.–3rd
ed.
p. cm.
ISBN-13 978-0-471-44917-1 (Cloth)
ISBN-10 0-471-44917-2 (Cloth)
1. Human engineering–Handbooks, manuals, etc. I. Salvendy, Gavriel, 1938-
TA166.H275 2005
620.8′2–dc22
2005003111

Printed in the United States of America

10 9 8 7 6 5 4 3 2

CHAPTER 27
HUMAN ERROR

Joseph Sharit
University of Miami
Coral Gables, Florida

1 INTRODUCTION

1.1 Does Human Error Exist?

Human error in human–system interaction was a major influence in establishing the area of human factors (Helander, 1997). Recognition of human error as an area in its own right came some years later, when an appreciation for its implications in complex high-risk systems became more widespread. Human error has since become inextricably linked to safety science, and together these areas have strongly influenced the design, reliability, risk assessment, and risk management programs that have become pivotal to the success of many organizations.

For a good part of the twentieth century the dominant perspective on human error by many U.S. industries was to attribute adverse outcomes to the persons whose actions were most closely associated to these events. We are now witnessing an almost complete turnaround in this perspective, to the extent that it has even become fashionable to reject the notion of human error. In this view the human is deemed to be a reasonable entity at the mercy of an array of design, organizational, and situational factors that can lead to behaviors external observers come to regard, although often unfairly, as human errors.

The appeal of this view should be readily apparent in each of the following two cases. The first case involves a worker who is subjected to performing a task in a restricted space. While attempting to reach for a tool, the worker's forearm inadvertently activates a switch, resulting in the emission of heat. Visual feedback concerning the activation is not possible, due to the awkward posture the worker must assume; tactile cues are not detectable due to requirements for wearing protective clothing; and although present, auditory feedback from the switch's activation is not audible, due to high noise levels. Residual vapors originating from a rarely performed procedure during the previous shift ignite, resulting in an explosion. In the second case, a worker adapts the relatively rigid and unrealistic procedural requirements dictated in a written work procedure to demands that continually materialize in the form of shifting objectives, constraints on resources, and changes in production schedules. Management tacitly condones these procedural adaptations, in effect relying on the resourcefulness of the worker for ensuring that its goals are met (Section 8). However, when an unanticipated scenario causes the worker's adaptations to result in an accident, management is swift to renounce any support of actions in violation of work procedures.

In the first case the worker's action that led to the accident was unintentional; in the second case the worker's actions were intentional. In both cases the issue of whether the "actor" committed an error is debatable. One variant on the position that rejects the notion of human error would shift the blame for the adverse consequences from the actor to management or the designers. Latent management or latent designer errors (Section 7.3.1) would thus absolve the actor from human error in each of these cases. The worker, after all, was in the heat of the battle, performing "normal work," responding to the contextual features of the situation in reasonable, even skillful ways.

A second variant on this position would cast doubt on the process by which the attribution of error is made (Dekker, 2005). By virtue of having knowledge of events, especially bad events such as accidents, outside observers are able—perhaps even motivated—to invoke a backward series of rationalizations and logical connections that has neatly filtered out the subtle and complex situational details that are likely to be the basis for the perpetrating actions. Whether this process of establishing causality (Section 10) is due to convenience, or derives from the inability to determine or comprehend the perceptions and assessments made by the actor that interlace the more prominently observable events, the end result is a considerable underestimation of the influence of context. Even the workers themselves, if given the opportunity in each of these cases to examine or reflect upon their performance, may acknowledge their actions as errors, easily spotting all the poor decisions and improperly executed actions, when in reality, within the frames of references at the time the behaviors occurred, their actions were in fact reasonable, and constituted "mostly normal work." The challenge, according to Dekker (2005), is "to understand how assessments and actions that from the outside look like errors become neutralized or normalized so that from the inside they appear unremarkable, routine, normal" (p. 75).

These views, which essentially deny the existence of human error (at least on the part of the actors) are appealing and to some extent justified. The issue, however, is not so much whether these views should be dismissed, but whether they should be embraced. The position taken here is that human error is a real phenomenon that has at its roots many of the same attentional processes and architectural features of memory that enable the human to adapt, abstract, infer, and create, but that also subject the human to various kinds of information-processing constraints that can provoke unintended or mistaken actions. Thus, although it may be convenient to explain unintended *action slips* (Section 3.2.5) such as the activation of an incorrect control or the selection of the wrong medication as rational responses in contexts characterized by pressures, conflicts, ambiguities, and fatigue, a closer inspection of the work context can, in theory, reveal the increased possibility for certain types of errors as compared to others. It is human fallibility, in all its guises, that infiltrates these contexts, and by failing to acknowledge the interplay between human fallibility and context—for instance, the tendency for a context to induce "capture" by the wrong control or the wrong medication—we are left with a shoddier picture of the context. Granted, the contextual details comprising dynamic work activities are difficult enough to establish, let alone their interplay with human fallibility. However, this fact attests only to the difficulty of predicting human error (Section 4), especially complex errors, not to the dismissal of its existence. Whereas rejecting the notion of human error may represent a gracious gesture toward the human's underlying disposition, it can also dangerously downplay aspects of human fallibility that need to be understood for implementing error reduction and error management strategies.

1.2 New Directions

Much of the practical knowledge that has been accumulated on human error in the last half century has derived primarily from industries requiring hazardous operations that are capable of producing catastrophic events. Not surprisingly, the textbook scenarios typically used for studying human error came from domains such as nuclear power, chemical processing, and aviation. With the publication of *To Err Is Human* (Kohn et al., 1999) came the revelation of shocking data that formally announced the new scourge in human error—medical error. According to this report, between 44,000 and 98,000 hospitalized patients die annually as a result of human error. These figures were extrapolated from studies that included the relatively well known Harvard Medical Practice Study in New York (Leape et al., 1991). Although these figures have been contested on the grounds that many of the patients whose deaths were attributed to medical error were predisposed to die due to the severity of their illnesses, there is also an opposing belief that these errors were underreported by as much as a factor of 10 (Cullen et al., 1995). If true, the number of preventable hospital deaths attributable to human error is staggering, even if adjustments are made for deaths that were likely to occur due to illness alone.

It also signals the need for heightened concern for the many mistakes in health care that are probably occurring outside hospital environments.

Fear of blame and retribution through litigation accounts for much of the underreporting in health care and reflects an industry that is still mired in the blame culture of traditional mid-twentieth-century American industry. What truly dissociates medical error from human error in other high-risk work domains is the belief by many people that they can assume the role of expert based on the experiences they, a family member, or a close friend have had, and that they have the right to hold an industry that is extracting high premiums for their services accountable for its actions. It remains to be seen if this attitude will carry over to other industries. In any case, medical error presents unique challenges, and although we do not intend to diminish the significance of human error in industries with relatively long-standing traditions for addressing the role of human error in safety, due emphasis will also be given to medical error.

2 DEFINING HUMAN ERROR

The presumption of human error generally occurs when various types of committed or omitted human actions appear, when viewed in retrospect, to be linked to undesirable consequences, although unwanted consequences do not necessarily imply the occurrence of human error. Following the distinctions proposed by Norman (1981) and Reason (1990), the term *error* usually applies only to those situations where there was an intention to perform some type of action, and would include cases where there was no prior intention to act. Thus, a very well practiced routine that is performed without any prior intention, such as swiping dirt from a tool, may constitute an error depending on the effect of that action. More typically, errors are associated with prior intentions to act, in which case two situations can be differentiated. If for whatever reason the actions did not proceed as planned, any unwanted consequences resulting from these actions would be attributed to an error arising from an unintentional action. In the case where the actions did proceed as intended but did not achieve their intended result, any unwanted outcomes stemming from these actions would be associated with an error resulting from intended but mistaken actions.

In each of these situations the common element is the occurrence of unwanted or adverse outcomes. Whether intended or not, negative outcomes need not be directly associated with these actions. Human error thus also subsumes actions whose unwanted outcomes may occur at much later points in time or following the interjection of many other actions by other people. It can also be argued that even if these actions did not result in adverse outcomes but had the potential to, they should be viewed as errors, in line with the current emphasis on near misses and the recognition that what separates many accidents from events with no visibly apparent negative consequences is chance alone. Acts of sabotage, although capable of bringing about adverse consequences, are not actions that deviate from expectations and thus do not constitute human error. Similarly, intentional violations of procedures, although also of great concern, are typically excluded from definitions of human error when the actions have gone as planned. For example, violations in rigid "ultrasafe and ultraregulated systems" are often required for effectively managing work constraints (Amalberti, 2001). However, when violations result in unforeseen and potentially hazardous conditions, these actions would constitute human error. Exploratory behavior under presumably protective or kind conditions as encountered in formal training programs or trial-and-error self-learning situations, which leads either to unintentional actions or mistaken actions should also be dissociated from human error. This distinction highlights the need to acknowledge the role of error—indeed, even the need for encouraging errors—in adaptation and creativity and in the acquisition of knowledge and skill associated with learning.

The situation becomes more blurred when humans knowingly implement strategies in performance that will result in some degree of error, as when a supervisor encourages workers to adopt shortcuts that trade off accuracy for speed, or when the human reasons that the effort needed to eliminate the possibility of some types of errors may increase the likelihood of more harmful errors. As with procedural violations, if these strategies come off as intended, the actor would not consider the attendant negative outcomes as having resulted from human error. However, depending on the boundaries of acceptable outcomes established or perceived by external observers such as managers or the public, the human's actions may in fact be considered to be in error. Accordingly, a person's ability to provide a reasonable argument for behaviors that resulted in unwanted consequences does not necessarily exonerate the person from having committed an error. What of actions the person intends to commit that are normally associated with acceptable outcomes but which result in adverse outcomes? These would generally not be considered to be human error except perhaps by unforgiving stakeholders who are compelled to exact blame.

The lack of consensus in arriving at a satisfying definition of human error is troubling in that it can undermine efforts to identify, control, and mitigate errors across different work domains and organizations. In fact, some authors have abandoned the term *human error* altogether. Hollnagel (1993) prefers the term *erroneous action* to human error, which he defines as "an action which fails to produce the expected result and which therefore leads to an unwanted consequence" (p. 67). Dekker's (2005) view of errors as "*ex post facto* constructs rather than as objective, observed facts" (p. 67) is based on the accumulated evidence on hindsight bias (Section 10.1). Specifically, the predisposition for this bias has repeatedly demonstrated how observers, including people who may have been recent participants of the experiences being investigated, impose their knowledge (in the form of assumptions and facts), past experiences, and future intentions

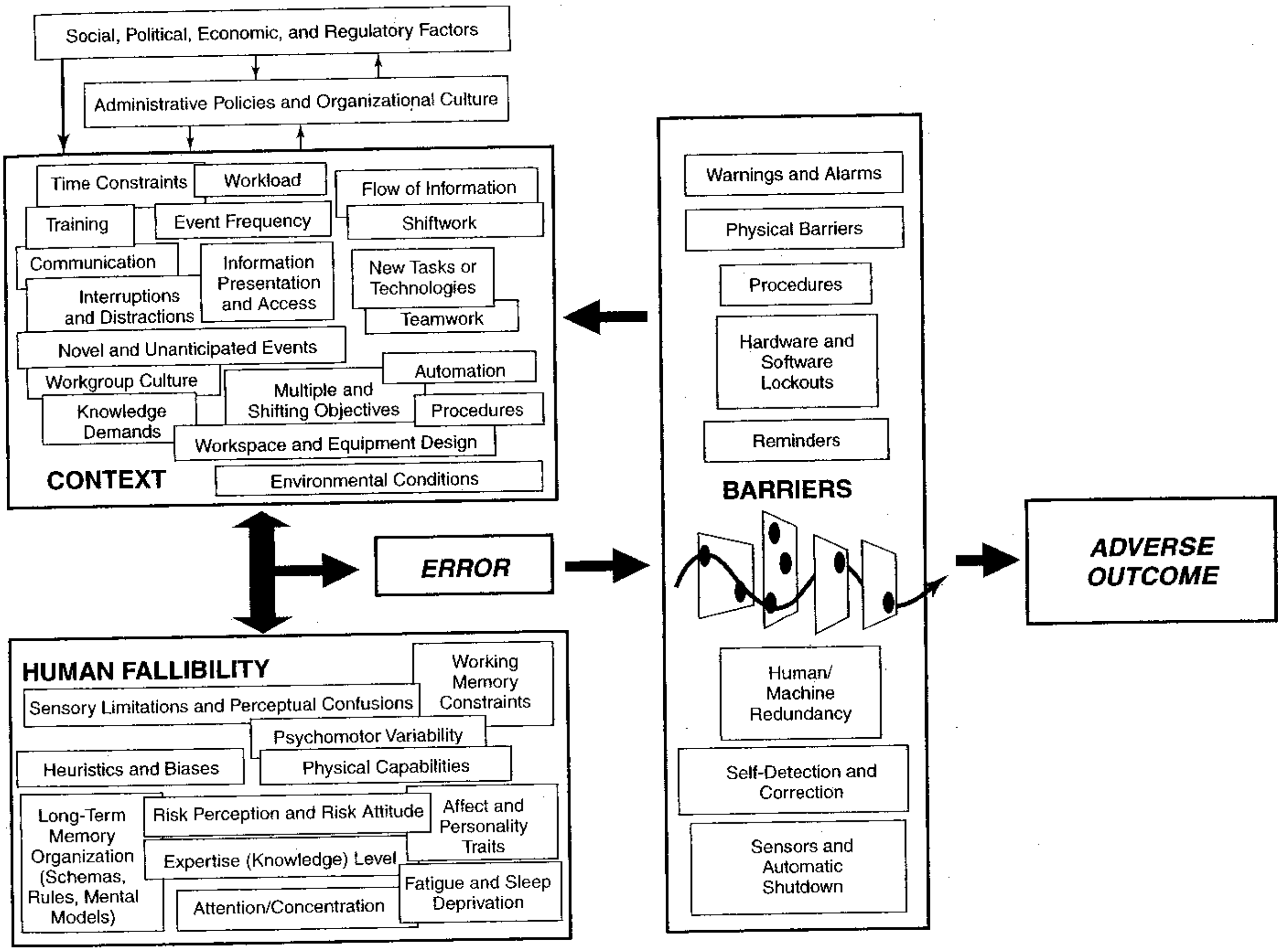

Social, Political, Economic, and Regulatory Factors
Administrative Policies and Organizational Culture
Time Constraints
Workload
Flow of Information
Training
Event Frequency
Shiftwork
Communication
Information Presentation and Access
New Tasks or Technologies
Interruptions and Distractions
Teamwork
Novel and Unanticipated Events
Automation
Workgroup Culture
Multiple and Shifting Objectives
Knowledge Demands
Procedures
Workspace and Equipment Design
CONTEXT
Environmental Conditions
Warnings and Alarms
Physical Barriers
Procedures
Hardware and Software Lockouts
Reminders
BARRIERS
ERROR
ADVERSE OUTCOME
HUMAN FALLIBILITY
Working Memory Constraints
Sensory Limitations and Perceptual Confusions
Psychomotor Variability
Heuristics and Biases
Physical Capabilities
Long-Term Memory Organization (Schemas, Rules, Mental Models)
Risk Perception and Risk Attitude
Affect and Personality Traits
Expertise (Knowledge) Level
Fatigue and Sleep Deprivation
Attention/Concentration
Human/ Machine Redundancy
Self-Detection and Correction
Sensors and Automatic Shutdown

Figure 1 Framework for understanding human error.

to transform what was in fact inaccessible information at the time into neatly unfolding sequences of events and deterministic schemes that are capable of explaining any adverse consequence. These observer and hindsight biases presumably do not bring us any closer to understanding the experiences of the actor in the actual situation for whom there is no error—"the error only exists by virtue of the observer and his or her position on the outside of the stream of experience" (p. 66).

Although this view is enlightening in its ability to draw attention to the limitations of empiricist-based paradigms that underlie many human factors methods, it is also subject to some of the same criticisms that were raised in Section 1.1 in response to the current trend toward perspectives that negate the existence of human error. Understanding both human fallibility and the contexts in which humans must act keeps us on a pragmatic path capable of shaping design and safety-related interventions, even as we strive to find methods that can close the gaps between objective and reconstructed experiences. As we shall see in Section 4, the problems associated with defining human error can be partly overcome by shifting the emphasis to classification schemes that are capable of establishing links between human psychological processes and the manifestation of adverse outcomes across different work domains.

3 UNDERSTANDING HUMAN ERROR

3.1 A Modeling Framework: Human Fallibility, Context, and Barriers

Figure 1 presents a simple modeling framework for demonstrating how human error arises and can result in adverse outcomes. There are three major components in this model. The first component, *human fallibility*, addresses the fundamental sensory, cognitive, and motor limitations of humans that predispose them to error. The second component, *context*, refers to situational variables that can affect the way in which human fallibility becomes manifest. The third component, *barriers*, concerns the various ways in which human errors can be contained.

A number of general observations concerning this modeling framework are worth noting. First, human error is viewed as arising from an interplay between human fallibility and context. This is probably the most intuitive way for practitioners to understand the causality of human error. Interventions that minimize human dispositions to fallibility, for example by placing fewer memory demands on the human, are helpful only to the extent that they do not create new contexts that can, in turn, create new opportunities for human fallibility to become manifest. Similarly, interventions intended to reduce the error-producing potential of work contexts, for instance, by introducing new protocols for communication, could unsuspectingly produce new ways in which human fallibility can exert itself. Second, the depiction of overlapping elements in the human fallibility and context components of the model (Figure 1) is intended to convey the interactive complexity that may exist among these factors. For example, memory constraints may result in the use of heuristics that, in certain contexts, may predispose the human to error; these same memory constraints may also produce misguided perceptions of risk likelihood. Similarly, training programs that dictate how work procedures should be implemented could lead to antagonistic work group cultures whose doctrines afford increased opportunities for operational errors.

Third, barriers capable of preventing the propagation of errors to adverse outcomes could also affect the context. This potential interplay between barriers and context is often ignored or misunderstood in evaluating a system's risk potential. Fourth, system states or conditions that result from errors can propagate into adverse outcomes such as accidents, but only if the gaps in existing barriers are aligned to expose such windows of opportunity (Reason, 1990). The likelihood that errors will penetrate these juxtaposed barriers, especially in high-risk work activities, is generally low and is the basis for the much larger number of near misses that are observed compared to events with serious consequences. Finally, this modeling framework is intended to encompass various perspectives on human error that have been proposed (CCPS, 1994)—in particular, the human factors and ergonomics, cognitive engineering, and sociotechnical perspectives.

In the human factors perspective, error is the result of a mismatch between task demands and human mental and physical capabilities. Presumably this perspective allows only general predictions of human error to be made—primarily predictions of errors that are based on their external characteristics. For example, cluttered displays or interfaces that impose heavy demands on working memory are likely to overload perceptual and memory processes (Section 3.2) and thus possibly lead to the omission of actions or the confusion of one control with another. Guidelines that have been proposed for designing displays (Wickens et al., 2004) are offered as a means for diminishing mismatches between demands and capabilities and thus the potential for error. In contrast, the cognitive engineering perspective emphasizes detailed analysis of work contexts (Section 4) coupled with analysis of the human's intentions and goals. Although both the human factors and cognitive engineering perspectives on human error are very concerned with human information processing, cognitive engineering approaches attempt to derive more detailed information about how humans acquire and represent information and how they use it to guide actions. This emphasis provides a stronger basis for linking underlying cognitive processes with the external form of the error, and thus should lead to more effective classifications of human performance and human errors. As a simple illustration of the cognitive engineering perspective, Table 1 demonstrates how the same external expression of an error could derive from various underlying causes.

Sociotechnical perspectives on human error focus on the potential impact of management policies and organizational culture on shaping the contexts within which people act. These "higher-order" contextual factors are capable of exacting considerable influence

Table 1 Examples of Different Underlying Causes of the Same External Error Mode

Situation: A worker in a chemical processing plant closes valve B instead of nearby valve A, which is the required action as set out in the procedures. Although there are many possible causes of this error, consider the following five possible explanations.

1. The valves were close together and badly labeled. The worker was not familiar with the valves and therefore chose the wrong one.
   Possible cause: wrong identification compounded by lack of familiarity leading to wrong intention (once the wrong identification had occurred the worker intended to close the wrong valve).

2. The worker may have misheard instructions issued by the supervisor and thought that valve B was the required valve.
   Possible cause: communications failure giving rise to a mistaken intention.

3. Because of the close proximity of the valves, even though he intended to close valve A, he inadvertently operated valve B when he reached for the valves.
   Possible cause: correct intention but wrong execution of action.

4. The worker closed valve B very frequently as part of his everyday job. The operation of A was embedded within a long sequence of other operations that were similar to those normally associated with valve B. The worker knew that he had to close A in this case, but he was distracted by a colleague and reverted back to the strong habit of operating B.
   Possible cause: intrusion of a strong habit due to external distraction (correct intention but wrong execution).

5. The worker believed that valve A had to be closed. However, it was believed by the workforce that despite the operating instructions, closing B had an effect similar to closing A and in fact produced less disruption to downstream production.
   Possible cause: violation as a result of mistaken information and an informal company culture to concentrate on production rather than safety goals (wrong intention).

Source: Adapted from CCPS (1994). Copyright 1994 by the American Institute of Chemical Engineers, and reproduced by permission of AIChE.

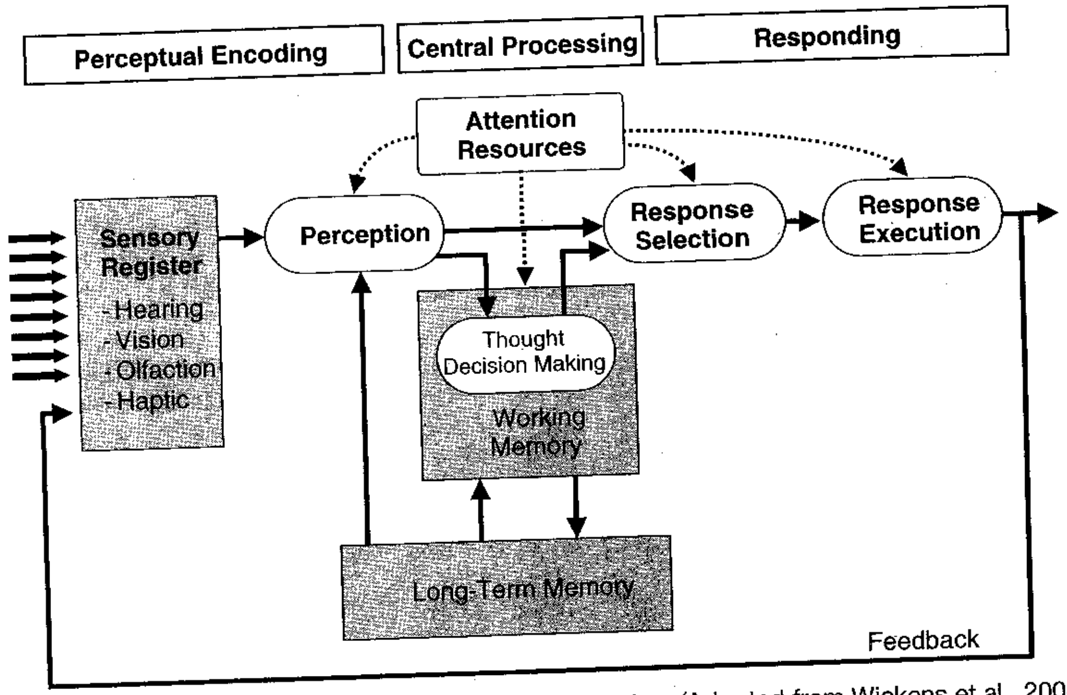


Figure 2 Generic model of human information processing. (Adapted from Wickens et al., 2004.)

on the designs of workplaces, operating procedures, training programs, job aids, and communication protocols, and can produce excessive workload demands by imposing multiple conflicting and shifting performance objectives and by exerting pressure to meet production goals, often at the expense of safety considerations. How work cultures (the cultures associated with those people responsible for producing products and services) become established is a complex phenomenon, and though numerous factors can play a role, the strongest influence is most likely to be the organizational climate (Section 9). Problematic work cultures are very resistant to change, and their remediation usually requires multiple interventions at both the management and operator levels over extended periods of time.

Although the human factors and ergonomics, cognitive engineering, and sociotechnical perspectives appear to suggest different approaches for predicting and analyzing human error, the study of human error will often require the collective consideration of these different perspectives. Human capabilities and limitations from a human factors and ergonomics perspective provide the fundamental basis for pursuing more rigorous cognitive engineering analyses of human error. Similarly, the cognitive engineering perspective, in its requirements for more detailed analyses of work contexts, would be remiss to exclude sociotechnical considerations (Chapter 10).

3.2 Human Fallibility

3.2.1 Human Information Processing

The basis for many human errors derives from fundamental limitations that exist in the human's sensory, cognitive, and motor processes (Chapter 5). These limitations are best understood by considering a generic model of human information processing (Wickens et al., 2004) that conceptualizes the existence of various *processing resources* for handling the flow and transformation of information (Figure 2).

According to this model, sensory information received by the body's various receptor cells gets stored in a system of sensory registers that has an enormous storage capacity. However, this information is available for further processing only briefly. Through the process of selective attention, subsets of this vast collection of information become designated for further processing in an early stage of information processing known as perception. Here, information can become meaningful through comparison with information in long-term memory (LTM), which may result in a response or the need for further processing in a short-term memory store referred to as working memory (WM). A good deal of our conscious effort is dedicated to WM activities such as visualizing, planning, evaluating, conceptualizing, and making decisions, and much of this WM activity depends on information that can be accessed from LTM. Rehearsal of information in WM enables it to be encoded into LTM; otherwise, it decays rapidly. WM also has relatively severe capacity constraints governing the amount of information that can be kept active. The current contention is that within WM there are separate storage systems for accommodating visual information in an analog spatial form or verbal information in an acoustical form, and an attentional control system for coordinating these two storage systems. Ultimately, the results of WM/LTM analysis can lead to a response (e.g., a motor action or decision), or to the revision of thoughts. Note that although this sequence of information processing is depicted in Figure 2 as flowing from left to right, in principle it can begin anywhere.

With the exception of the system of sensory registers and LTM the processing resources in this model may require *attention*. Often thought of as mental effort, attention is conceptualized here as a finite and flexible internal energy source under conscious control whose intensity can be modulated over time. Although attention can be distributed among the information-processing resources, fundamental limitations in attention constrain the capacities of these resources—that is, there is only so much information that can undergo perceptual coding or WM analysis. Focusing attention on one of these resources will, in many cases, handicap other resources. Thus if a North American rents a car with a manual transmission in Great Britain, the experience of driving on the left-side of the road may require substantial allocation of attention to perceptual processing in order to avoid collisions with other drivers, perhaps at the expense of being able to smoothly navigate the stick shift (which is now located to the left of the driver), or at the expense of using WM resources to keep adequate track of one's route. Whatever attention is allocated to WM may be needed for working out the cognitive spatial transformations required for executing left-hand and right-hand turns.

Attention may also be focused almost exclusively on WM, as often occurs during intense problem solving or when planning activities. The ability to divide attention, which is the basis for time sharing, is often observed in people who may have learned to rapidly shift attention between tasks. This skill may require knowledge of the temporal and knowledge demands of the tasks and the possibility for one or more of the tasks having become automated in the sense that very little attention is needed for their performance. Various dichotomies within the information-processing system have been proposed, for example, between the visual and auditory modalities and between early (perceptual) versus later (central and response) processing (Figure 2), to account for how people are able, in time-sharing situations, to more effectively utilize their processing capacities (Wickens, 1984).

Many design implications arise from the errors that human sensory and motor limitations can cause or contribute to. Indeed, human factors studies are often preoccupied with deriving design guidelines for minimizing such errors. Knowledge concerning human limitations in contrast sensitivity, hearing, bandwidth in motor movement, and in sensing tactile feedback can be used to design visual displays, auditory alarms, manual control systems, and protective clothing (such as gloves that are worn in surgery) that are less likely to produce errors in detection and response. Much of the focus on human error, however, is on the role that cognitive processing plays. Even seemingly simple situations involving errors in visual processing may in fact be rooted in much more complex information processing as illustrated in the following example.

3.2.2 Example: Medication Error

Consider the following prescription medication error, which actually occurred. A physician opted to change the order for 50 mg of a leukemia drug to 25 mg by putting a line through the zero and inserting a "2" in front of the "5." The resulting dose was perceived by the pharmacist as 250 mg and led to the death of a 14-year-old boy. The line that was meant to indicate a cross-out was not centered and turned out to be much closer to the right side of the circle (due to *psychomotor variability*; see Figure 1); thus, it could easily have been construed as just a badly written zero. Also, when one considers that perception relies on both bottom-up processing (where the stimulus pattern is decomposed into features) and top-down processing (where context and thus expectations are used for recognition), the possibility that a digit was crossed out may have countered expectations (i.e., it does not usually occur).

If one were to presume that the pharmacist had a high workload (and thus diminished resources for processing the prescription) and a relative lack of experience or knowledge concerning dosage ranges for this drug, it is easy to understand how this error can come about. The dynamics of the error can be put into a more complete perspective when potential barriers are considered, such as an automatic checking system that could have screened the order for a potentially harmful dosage or interactions with other drugs, or a procedure that would have required the physician to rewrite any order that had been altered. Even if these barriers were in place, which was not the case, there is a high likelihood that they would be bypassed. In fact, if such a procedure were to be imposed on physicians, routine violations would be expected given the contexts within which many physicians work.

3.2.3 Long-Term Memory and Its Implications for Human Error

LTM has been described as a parallel distributed architecture that is being reconfigured continuously through selective activation and inhibition of massively interconnected neuronal units (Rumelhart and McClelland, 1986). These reconfiguration processes occur within distinct modules that are responsible for different representations of information, such as mental images or sentence syntax. In the process of adapting to new stimuli or thoughts, the complex interactions between neuronal units that are produced give rise to the generalizations and rules that are so critical to human performance. With regard to the forms of knowledge stored in LTM, we usually distinguish between the general knowledge we have about the world, referred to as *semantic memory*, and knowledge about events, referred to as *episodic memory*.

When items of information, such as visual images, sounds, and thoughts based on existing knowledge, are processed in WM at the same time, they become associated with each other in LTM. The retrieval of this information from LTM will then depend on the strength of the individual items as well as the strengths of their associations with other items. Increased frequency and recency of activation are assumed to promote stronger (i.e., more stable) memory traces, which are otherwise subject to negative exponential decays.

Much of our basic knowledge about things can be thought of as being stored in the form of *semantic networks* that are implemented through parallel distributed architectures. Other knowledge representation schemes commonly invoked in the human factors literature are schemas and mental models. *Schemas* typically represent knowledge organized about a concept or topic. When they reflect processes or systems for which there are relationships between inputs and outputs that the human can mentally visualize and experiment with (i.e., "run," like a simulation program), the schemas are often referred to as *mental models*. The organization of knowledge in LTM as schemas or mental models is also likely based on semantic networks.

The constraints associated with LTM architecture can provide many insights into human fallibility and how this fallibility can interact with situational contexts to produce errors. For example, implicit in the existence of parallel associative networks is the ability to recall both items of information and patterns (i.e., associations) of information based on partial matching of this information with the contents of memory. Because the contexts within which humans operate often produce what Reason (1990) has termed *cognitive underspecification*, the implication is that at some point in the processing of information the specification of information may be incomplete. It may be incomplete due to perceptual processing constraints, WM constraints, or LTM (i.e., knowledge) limitations, or due to external constraints, as when there is little information available on the medical history of a patient undergoing emergency treatment or when piping and instrumentation diagrams have not been updated. LTM organization can overcome these limitations by retrieving some items of information that provide a match to the inputs, and thus enable an entire rule, by previous association with other items of information in LTM, to be activated. Unfortunately, that rule may not be appropriate for the particular situation. Similarly, for instance in the case of a radiologist who has recently encountered a large number of tumors of a particular type, the increased activation levels that are likely to be associated with this diagnosis may result in a greater tendency for arriving at this diagnosis in future situations.

3.2.4 Information Processing and Decision-Making Errors

Human decision making that is not guided by normative prescriptive models (Chapter 8) is an activity fraught with fallibility, especially in complex dynamic environments. As illustrated in Figure 3, human limitations in decision making can arise from a number of information-processing considerations (Figure 2) that directly or indirectly implicate LTM (Wickens et al., 2004). For example, if the information the human opts to select for WM activity, which may be guided by past experiences, is fuzzy or incomplete, intensive interpretation or integration of this information may be needed. Also, any hypotheses that the decision maker generates regarding this information will be highly dependent on information that can be retrieved from LTM, and their evaluation could require searching for additional information. Although any hypothesis for which adequate support is found can become the basis for an action, the possible candidate actions that would need to be evaluated in WM would first need to be retrieved from LTM. In addition, the possible outcomes associated with each action, the estimates of the likelihoods of

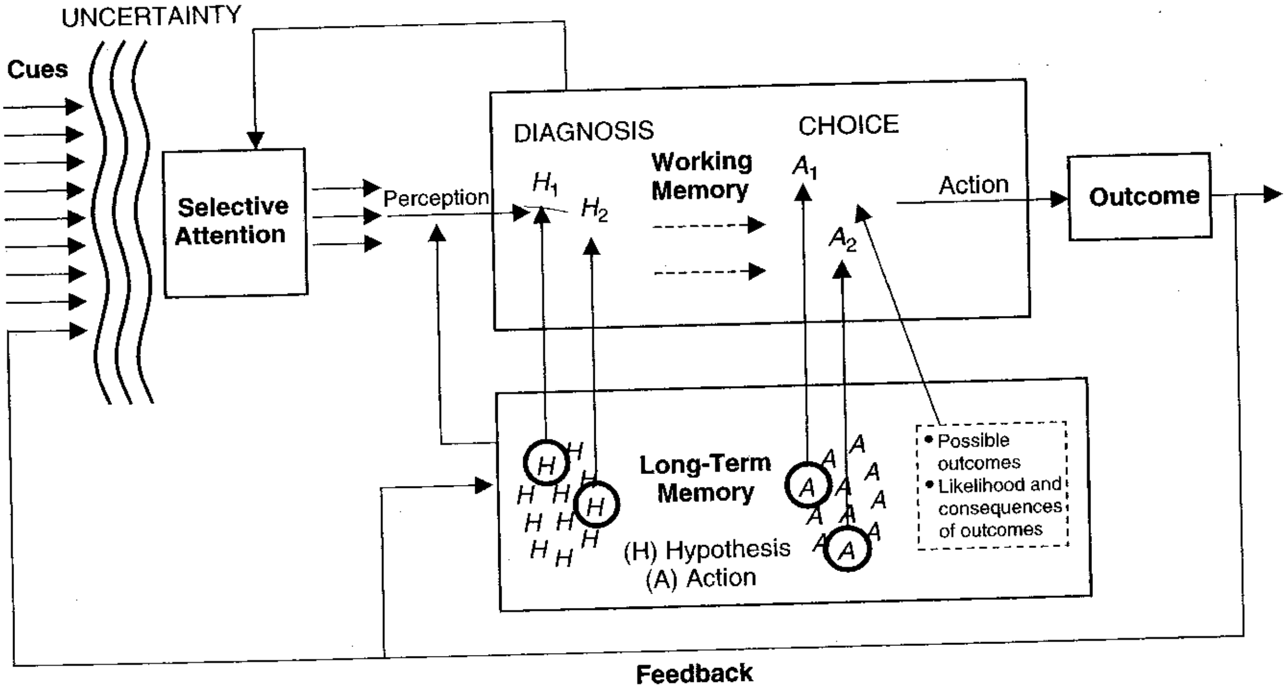


Figure 3 Information-processing model of decision making. (Adapted from Wickens et al., 2004.)

these outcomes, and the negative and positive implications of these actions would also require retrieval from LTM.

From an information-processing perspective, there are numerous factors that could constrain this decision-making process, particularly factors that could influence the amount or quality of information brought into WM and the retrieval of information from LTM. These constraints often lead to shortcuts in decision making, such as *satisficing* (Simon, 1966), whereby people opt for choices that are good enough for their purposes and adopt strategies for sampling information that they perceive to be most relevant. In general, the human's natural tendency to minimize cognitive effort (Sharit, 2003) opens the door to a wide variety of shortcuts or heuristics that are efficient and usually effective in negotiating environmental complexity, but under the right coincidence of circumstances can lead to ineffective choices or actions that become designated as errors. For example, with respect to the cues of information that we perceive, there is a tendency to overweight cues occurring earlier than later in time or that change over time. WM will only allow for a limited number of possible hypotheses, actions, or outcomes of actions to be evaluated, and LTM architecture will accommodate these limitations by making information that has been considered more frequently or recently (the "availability" heuristic) more readily available and by enabling its partial-matching capabilities to classify cues as more representative of a hypothesis than may be warranted. Many other heuristics (Wickens et al., 2004), such as *confirmation bias* (the tendency to consider confirming and not disconfirming evidence when evaluating hypotheses), *cognitive fixation* (remaining fixated on initial hypotheses and underutilizing subsequent information), and the tendency to judge an event as likely if its features are representative of its category (e.g., judging a person as having a particular occupation based on the person's appearance even though the likelihood of having that occupation is extremely low) derive primarily from a conservation of cognitive effort.

An enormous investment by the human in WM activities (i.e., an extensive commitment to functioning in an attentional mode) would be required to expose the biases that these heuristics can potentially induce. It is important to note, however, that to exclude the possibility that a human's situational assessments are in fact rational, explanations of human judgments and behaviors on the basis of cognitive biases require a sound understanding of the specific context (Fraser et al., 1992).

3.2.5 Levels of Human Performance and Dispositions for Errors

Considerations related to LTM architecture enable many different types of human errors to be accounted for by a few powerful principles. This few-to-many mapping between underlying memory mechanisms and different error types will be influenced by the nature of human performance, particularly on how information-processing resources (Figure 2) are used, other aspects of human fallibility (Section 3.2.6), and situational variables. A framework that distinguishes between skill-based, rule-based, and knowledge-based levels of performance—Rasmussen's *SRK framework*—emphasizes fundamentally different approaches to processing information and is thus particularly appealing for understanding the role of human performance in analyzing and predicting different types of human errors (Rasmussen, 1986).

Activities performed at the skill-based level are highly practiced routines that require little conscious attention. Referring to Figure 2, these activities map perception directly to actions, bypassing WM. Following an intention for action that could originate in WM or from environmental cues, the responses associated with the intended activity are so well integrated with the activity's sensory features that they are elicited in the form of highly automatic routines. Given the frequent repetitions of consistent mappings from sensory features to motor responses, the meaning imposed on perception by LTM can be thought of as hardwired to the human's motor response system.

The rule-based level of performance makes use of rules that have been established in LTM based on past experiences. WM is now a factor, as rules (of the if–then type) or schemas may be brought into play following the assessment of a situation or problem. More attention is thus required at this level of performance, and the partial matching characteristics of LTM can prove critical. When stored rules are not effective, as is often the case when new or challenging problems arise, the human is usually forced to devise plans that involve exploring and testing hypotheses, and must continuously refine the results of these efforts into a mental model or representation that can provide a satisfactory solution. At this knowledge-based level of performance heavy demands on information-processing resources are exacted, especially on WM, and performance is vulnerable to LTM architectural constraints to the extent that WM is dependent on LTM for problem solving.

In reality, many of the meaningful tasks that people perform represent mixtures of skill, rule, and knowledge-based levels of performance. Although performance at the skill-based level results in a significant economy in cognitive effort, the reduction in resources of attention comes at a risk. For example, consider a task other than the one that is intended that contains features that are similar to those of the intended task. If the alternative activity is frequently performed and therefore associated with skill-based automatic response patterns, all that is needed is a context that can distract the human from the intention and allow the human to be "captured" by the alternative (incorrect) task. This situation represents example 4 in Table 1 in the case of an inadvertent closure of a valve. In other situations the capture by a skill-based routine may result in the exclusion of an activity. For example, suppose that task A is performed infrequently and task B is performed routinely at the skill-based level. If the initial steps are identical for both tasks but task A requires an additional step, this step is likely to be omitted during execution of the task. Untimely interruptions are often the basis for omissions at the skill-based level of performance. In some circumstances, interruptions or moments of inactivity during skill-based routines may instigate thinking about where one is in the sequence of steps. By directing attention to routines that are not designed to be examined, steps could be performed out of sequence (reversal errors) or be repeated (Reason, 1990).

Many of the errors that occur at the rule-based level involve inappropriate matching of either external cues or internally generated information with the conditional components of rules stored in LTM. Generally, conditional components of rules that have been satisfied on a frequent basis or that appear to closely match prevailing conditions are more likely to be activated. The prediction of errors at this level of performance will thus require knowing what other rules the human might consider, thus necessitating detailed knowledge not only about the task but also about the process (e.g., training or experience) by which the person acquired rule-based knowledge. Mistakes in applying rules generally involve the misapplication of rules with proven success or the application of *bad rules* (Reason, 1990). Mistakes in applying rules with proven success often occur when *first exceptions* are encountered. Consider the case of an endoscopist who relies on indirect visual information when performing a colonoscopy. Based on past experiences and available knowledge, the sighting of an anatomical landmark during the performance of this procedure may be interpreted to mean that the instrument is situated at a particular location within the colon, when in fact the presence of an anatomical deformity in this patient may render the physician's interpretation as incorrect (Cao and Milgram, 2000). These first exception errors often result in the decomposition of general rules into more specific rule forms and reflect the acquisition of expertise. General rules, however, usually have higher activation levels in LTM given their increased likelihood of encounter, and under contextual conditions involving high workload and time constraints, they will be the ones more likely to be invoked. Rule-based mistakes that occur by applying bad (e.g., inadvisable) rules are also not uncommon, as when a person who is motivated to achieve high production values associates particular work conditions with the opportunity for implementing shortcuts in operations.

At the knowledge-based level of performance, when needed associations or schemas are not available in LTM, control shifts primarily to intensive WM activities. This level of performance is often associated with large degrees of freedom that characterize how a human "moves through the problem space," and suggests a much greater repertory of behavioral responses and corresponding expressions of error. Contextual factors that include task characteristics and personal factors that include emotional state, risk attitude, and confidence in intuitive abilities can play a significant role in shaping the error modes, making these types of errors much harder to predict. It is at this level of performance that we observe undue weights given to perceptually salient cues or early data, confirmation bias, use of the availability and representative heuristics (especially for assessing relationships between causes and effects), underestimation and overestimation of the likelihood of events in response to observed data, vagabonding (darting from issue to issue, often not even realizing that issues are being revisited, with essentially no effective movement through the problem space), and encysting (overattention to a few details at the expense of other, perhaps more relevant information).

3.2.6 Other Aspects of Human Fallibility

There are many facets to human fallibility, and all have the potential to contribute to human error. For example, personality traits that reflect dispositions toward confidence, conscientiousness, and perseverance could influence both the possibility for errors and the nature of their expression at both the rule- and knowledge-based levels of performance, especially under stress. Overconfidence can lead to risk-taking behaviors and has been implicated as a contributory factor in a number of accidents.

Sleep deprivation and fatigue are forms of human fallibility whose manifestations are often regarded as contextual factors. In fact, in the maritime and commercial aviation industries, these conditions are often attributed to company or regulatory agency rules governing hours of operation and rest time. The effects of fatigue may be to regress skilled performers to the level of unskilled performers (CCPS, 1994) through widespread degradation of abilities that include decision making and judgment, memory, reaction time, and vigilance. NASA has determined that about 20% of incidents reported to its Aviation Safety Reporting System (Section 6.3), which asks pilots to report problems anonymously, are fatigue-related (Kaye, 1999a). On numerous occasions pilots have been found to fall asleep at the controls, although they usually wake up in time to make the landing.

Another facet of human fallibility with important implications for human error is *situation awareness* (Chapter 20), which refers to a person's understanding or mental model of the immediate environment (Endsley, 1995). As in the case of fatigue, situation awareness represents an aspect of human fallibility that can be heavily influenced by contextual factors. In principle, any factor that could disrupt a human's ability to acquire or perceive relevant data concerning the elements in the environment, or compromise one's ability to understand the importance of that data and relate the data to events that may be unfolding in the near future, presumably can degrade situation awareness. Comprehending the importance of the various types of information in the environment also implies the need for temporal awareness—the need to be aware of how much time tasks require and how much time is available for their performance (Grosjean and

Terrier, 1999). Thus, potentially many factors related to both human fallibility and context can influence situation awareness. Increased knowledge or expertise should allow for better overall assessments of situations, especially under conditions of high workload and time constraints, by enabling elements of the problem and their relationships to be identified and considered in ways that would be difficult for those who are less familiar with the problem. In contrast, poor display designs that make integration of data difficult can easily impair the process of assessing situations. In operations involving teamwork, situation awareness can become disrupted by virtue of the confusion created by the presence of too many persons being involved in activities.

Finally, numerous affective factors can corrupt a human's information-processing capabilities and thereby predispose the human to error. Personal crises could lead to distractions, and emotionally loaded information can lead to the substitution of relevant information with "information trash." Similarly, a human's susceptibility to panic reactions and fear can impair information-processing activities critical to human performance.

3.3 Context

Human actions are embedded in contexts and can only be described meaningfully in reference to the details of the context that accompanied and produced them (Dekker, 2005). The possibility for human fallibility to result in human error as well as the expression of that error will thus depend on the context in which task activities occur. Although the notion of a context is often taken as obvious, it is not easy to define, leading to commonly encountered alternative expressions, such as *scenario*, *situation*, *situational context*, *situational details*, *contextual features*, *contextual dynamics*, *contextual factors*, and *work context*. Designers of advanced computing applications often speak in terms of providing functionalities that are responsive to various user contexts. Building on a definition of context proposed by Dey (2001) in the domain of context-aware computer applications, *context* is defined as any information that can be used to characterize the situation of a person, place, or object, as well as the dynamic interactions among these entities. This definition of context would regard a process such as training as an entity derived from these interactions and would also encompass information concerning how situations are developing and the human's responses to these situations.

Figure 1 reveals some representative contextual factors. In this depiction, the presumption is that higher-order *context-shaping factors* can influence contextual factors that are more directly linked to human performance. Contexts ultimately derive from the characterization of these factors and their interactions. Analysis of the interplay of human fallibility and context as a basis for understanding human error will be beneficial to the extent that relevant contextual factors can be identified and analyzed in detail.

A number of quantitative approaches to human error assessment (Section 5) employ concepts that are related to context. For example, several of these approaches use *performance-shaping factors* (PSFs) to either modify the probability estimate assigned to an activity performed in error (Swain and Guttmann, 1983) or as the basis for the estimation of human error (Embrey et al., 1984). Any environmental, individual, organizational, or task-related factor that could influence human performance can, in principle, qualify as a PSF; thus PSFs appear to be related to contextual factors. These approaches, however, by virtue of emphasizing probabilities as opposed to possibilities for error, assume additive effects of PSFs on human performance rather than interactive effects. In contrast, implicit to the concept of a context is the interactive complexity among contextual factors. A sociotechnical method for quantifying human error referred to as *STAHR* (Phillips et al., 1990) is somewhat more consistent with the concept of context than approaches based on PSFs. This method utilizes a hierarchical network of influence diagrams to represent the effects of direct influences on human error, such as time pressure and quality of training, as well as the effects of less direct influences, such as organizational and policy issues, which project their influences through the more direct factors. However, while STAHR imposes a hierarchical constraint on influences, the concept of context implicit to Figure 1 imposes no such constraint, thus enabling influences to be represented as an unconstrained network (Figure 4).

Generally, the emphasis on predicting the possibility for error as opposed to the probability of error relaxes the assessments required of contextual factors. In making these assessments, some of the possible considerations could include the extent to which a contextual factor is present (i.e., the level of activation of a network node) and the extent to which it can influence other factors (i.e., the level of activation of a network arc), as illustrated in Figure 4. Temporal characteristics underlying these influences could also be included. Also, as conceptualized in Figure 1, contextual factors can be refined to any degree of detail, and practitioners and analysts would need to determine for specific task domains of interest the appropriate level of contextual analysis. For example, the introduction of new technology into activities involving teamwork (Section 7.3) would require the characterization of each person's role with respect to the technology as well as analysis of how team communication may become altered as a consequence of these new roles. Links to other contextual factors come to mind immediately. The creation of new tasks may result in fragmented jobs that impose higher workload demands and less reliable mental models, due to the difficulty in forming meaningful associations in memory. These factors, in turn, can affect adversely communication among team members. New training protocols that do not anticipate many of these influences may further predispose the human to error by directing attention away from important cues.

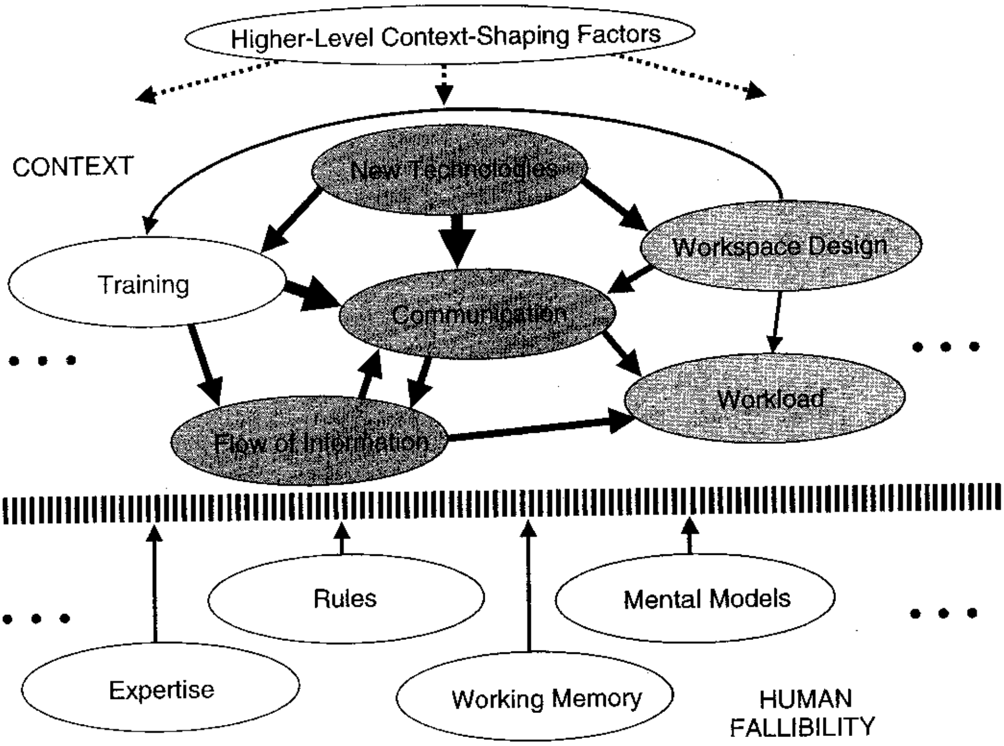


Figure 4 Influences between contextual factors representing part of a relevant work context, and their potential interplay with representative human fallibility factors. Activation levels of contextual factors are denoted by different degrees of shading and their degrees of influence are denoted by arrow widths. Temporal characteristics associated with these influences could also be included. Human fallibility factors can affect contextual factors as well as their influences. Influences among human fallibility factors are not depicted.

In some models of accident causation, the concept of a triggering event is used to draw attention to a "spark," such as a distraction or a pipe break, which sets off a chain of events (including human errors) that can ultimately lead to an accident. As defined here, a trigger represents just another contextual factor. Some triggering events, such as the random failure of a pump, may not necessarily have any contextual factors influencing it, whereas other triggers (e.g., a disruption in a work process resulting from a late discovery that a needed tool is absent) could very well be influenced by other contextual factors. Investigations of accidents (Section 10) are often aimed at exposing multiple chains of causal events and identifying critical paths whose disruption could have prevented the accident. Similarly, there may be nodes or paths within the network of contextual factors (Figure 4) that may be more responsive to human fallibility, and closer examination of these nodes and their links may inform the analyst of strategies for reducing human error or adverse events.

Finally, the possibility also exists for describing larger-scale work domain contexts that are capable of bringing about adverse outcomes through their interplay with human fallibility. In this regard, the views of Perrow (1999), which constitute a *system theory of accidents*, have received considerable attention. According to Perrow, the structural analysis of any system, whether technological, social, or political, reveals two loosely related concepts or dimensions, interactive complexity and coupling, whose sets of attributes govern the potential for adverse consequences. *Interactive complexity* can be categorized as either *complex* or *linear* and applies to all possible system components, including people, materials, procedures, equipment, design, and the environment. The relatively stronger presence of features such as reduced proximity in the spacing of system components, increased interconnectivity of subsystems, the potential for unintended or unfamiliar feedback loops, the existence of multiple and interacting controls (which can be administrative as well as technological), the presence of information that tends to be more indirect and incomplete, and the inability to easily substitute people in task activities predispose systems toward being complex as opposed to linear. Complex interactions are more likely to be produced by complex systems than linear systems, and because these interactions tend to be less perceptible and comprehensible the human's responses to problems that occur in complex systems can often further increase the system's interactive complexity.

Most systems can also be characterized by their degree of coupling. Tightly coupled systems are much less tolerant of delays in system processes than are loosely coupled systems and are much more invariant to materials and operational sequences. Although each type of system has both advantages and disadvantages, loosely coupled systems provide more opportunities for recovery from events with potentially adverse consequences, often through creative, flexible, and adaptive responses by people. To compensate for the fewer opportunities for recovery provided by tightly coupled systems, these systems generally require more built-in safety devices and redundancy than do loosely coupled systems.

Although Perrow's account of technological disasters focuses on the properties of systems themselves rather than human error associated with design, operation, or management of these systems, many of the catastrophic accidents chronicled by Perrow do in fact concern interactions between technological, human factors, organizational, and sociocultural systems, and technical systems are in their own right economic, social, and political constructs. Thus, despite the virtue in his theory of dispelling such accidents as having resulted from human error, his model has been criticized for its marginalization of factors at the root of technological accidents (Evan and Manion, 2002). These criticisms, however, do not preclude the possibility of augmenting Perrow's model with additional perspectives on system processes that would endow it with the capability for providing a reasonably compelling basis for predisposing the human to error.

3.4 Barriers

Various methods exist for building in barriers to human error. For example, computer-interactive systems can force the user to correct an invalid entry prior to proceeding, provide warnings about actions that are potentially error inducing, and employ self-correction algorithms that attempt to infer the user's intentions. Unfortunately, each of these methods can also be breached, depending on the context in which it is used. Forcing functions can initiate a process of backtracking by the user that can lead to total confusion and thus more opportunity for error (Reason, 1990), and warnings can be ignored under high workloads.

The facilitation of errors by computer-interactive systems was found to occur in a study by Koppel et al. (2005) on the use of hospital-computerized physician order-entry (CPOE) systems, contradicting widely held views that these systems significantly reduce medication prescribing errors. In this study, errors were grouped into two categories: (1) information errors arising from the fragmentation of data and the failure to integrate information across the various hospital information systems, and (2) human–machine interface flaws that fail to adequately consider the practitioner's behaviors in response to the constraints of the hospital's organizational work structure. An example of an error related to the first category is when the physician orders new medications or modifies existing medications. If current doses are not first discontinued, the medications may actually become increased or decreased, or be added on as duplicative or conflicting medication. Detection of these errors is hindered by flaws in the interface that may require 20 screens for viewing a single patient's medications. Complex organizational systems such as hospitals can make it extremely difficult for designers to anticipate the many contexts and associated problems that can arise from interactions with the systems that they design (Section 7.3.1). Although it may make more sense to have systems such as CPOEs monitored by practitioners and other workers for their error-inducing potential rather than have designers attempt to anticipate all the contexts associated with the use of these systems, this imposes the added burden of ensuring that mechanisms are in place for collecting the appropriate data, communicating this information to designers, and validating that the appropriate interventions have been incorporated.

Many of the electronic information devices (including CPOEs) that are currently in use in complex systems such as health care were implemented under the assumption that they would decrease the likelihood of human error. However, the benefits of reducing or even eliminating the possibility for certain types of errors often come at the risk of new errors, exemplifying how the introduction of barriers can create new windows of opportunity for errors through the alteration of existing contexts (Section 3.1). For example, in hospital systems the reliance on information in electronic form can disturb critical communication flows and is less likely than face-to-face communication to provide the cues and other information necessary for constructing appropriate models of patient problems.

One of the most frequently used barriers in industry—the written work procedure—is also one that is highly vulnerable to violation. Many of the procedures designed for high-hazard operations include warnings, contingencies (information on when and how to "back out" when dangerous conditions arise during operations), and other supporting features. To avoid the recurrence of past incidents, these procedures are updated continuously. Consequently, they grow in size and complexity to the point where they can contribute to information overload, increasing the possibility of missing or confusing important information (Reason, 1997). Procedures that disrupt the momentum of human actions are especially vulnerable to violation.

Humans themselves are quite adept at detecting and correcting many of the skill-based errors they make and are thus often relied upon to serve as barriers. Self-correction, however, implies two conditions: that the human depart from automated processing, even if only momentarily, and that the human invest attentional resources periodically to check whether the intentions are being met and that cues are available to alert one to deviation from intention (Reason, 1990). This would apply to both slips and omissions of actions. Redundancy in the form of cues presented in multiple modalities is a simple and very effective way of increasing a person's likelihood of detecting and correcting these types of errors. This strategy is illustrated in the case of the ampoule-swap error in hospital operating rooms (Levy et al., 2000). Many drug solutions are contained in ampoules that do not vary much in size and shape, often contain clear liquid solutions, and have few distinguishing features. If an anesthesiologist uses the wrong ampoule to fill a syringe and inadvertently "swaps in" a risky drug such as potassium chloride, serious consequences could ensue. Contextual factors such as fatigue and distractions make it unreasonable to expect medical providers to invest the attentional resources necessary for averting these types of errors. Moreover, the use of warning signs on bins that store ampoules containing "risky solutions" are poor solutions to this problem, as they require that the human maintain *knowledge in the head*—specifically, in WM—thus making this information vulnerable to memory loss resulting from delays or distractions between retrieving the ampoule and preparing the solution. The more reliable solution that was suggested by these investigators was to provide tactile cues on both the storage bins and the ampoules. For example, wrapping a rubber band around the ampoule following its removal from the bin provides an alerting cue in the form of tactile feedback prior to loading the ampoule into the syringe.

Not surprisingly, the human's error detection abilities are greatly reduced at the knowledge-based level of performance. Error detection in these more complex situations will depend on discovering that the wrong goal has been selected or recognizing that one's movement through the problem space is not consistent with the goal. In this regard, strategic errors (e.g., in goal definition) are expected to be much harder to discover than tactical errors (e.g., in choosing which subsystem to diagnose). Human error detection and recovery at the knowledge-based level of performance may in fact represent a highly evolved form of expertise. Interestingly, whereas knowledge-based errors decrease with increased expertise, skill-based errors increase. Also, experienced workers, as compared to beginners, tend to disregard a larger number of errors that have no work-related consequences, suggesting that with expertise comes the ability to apply higher-order criteria for regulating the work system, thus enabling the allocation of attention to errors to occur on a more selective basis (Amalberti, 2001).

A very common barrier to human error is having other people available for error detection. As with hardware components, human redundancy will usually lead to more reliable systems. However, successful human redundancy often requires that the other people be external to the operational situation, and thus possibly less subject to tendencies such as cognitive fixation. In a study of 99 simulated emergency scenarios involving nuclear power plant crews, Woods (1984) found that none of the errors involving diagnosis of the system state were detected by the operators who made them and that only other people were able to detect a number of them. In contrast, half the errors categorized as slips (i.e., errors in execution of correct intentions) were detected by the operators who made them. These results also suggest that team members can often be subject to the same error-producing tendencies as individuals.

Barriers to human error need not always be present by design. As implied in Perrow's system theory of accidents (Section 3.3), a complex mixture of a system's properties can produce conditions that are conducive to human error as well as to its detection and correction. This phenomenon is routinely demonstrated in large-scale hospital systems where one encounters an assortment of patient problem scenarios, a variety of health care services, complex flows of patient information across various media on a continual 24-hour basis, and a large variability in the skill levels of health care providers, who must often perform under conditions of overload and fatigue while being subjected to various administrative constraints. The complex interactions that arise under these circumstances provide multiple opportunities for human error, arising from missed or misunderstood information or confusion in following treatment protocols. Fortunately, there usually exist multiple layers of redundancy in the form of alternative materials (e.g., medications and equipment), treatment schedules, and health care workers to thwart the serious propagation of many of these errors. Thus, despite a number of constraints that exist in hospital systems, particularly in the provision of critical care, these systems are sufficiently loosely coupled to overcome many of the risks that arise in patient care, including those that are generated by virtue of discontinuities or gaps in treatment (Cook et al., 2000). However, even if adverse consequences are indeed averted in many of these cases, one must acknowledge the possibility that the quality of patient care may become significantly compromised in the process.

Finally, there is always the possibility that the perceived presence of barriers such as intelligent sensing systems and corrective devices may actually increase a person's risk-taking behavior. Adjusting risk-taking behavior to maintain a constant level of risk is in line with *risk-homeostasis theory* (Wilde, 1982). These adjustments presume that humans are reasonably good at estimating the magnitude of risk, which generally does not appear to be the case. Nonetheless, a disturbing implication of this theory is the possibility that interventions by organizations directed at improving the safety climate could, instead, result in work cultures that promote attitudes that are not conducive to safe operations.

3.5 Example: Wrong-Site Surgery

Wrong-site errors in health care encompass surgical procedures performed on a wrong part of the body, wrong side of the body, wrong person, or at the wrong level of a correctly identified anatomical site. The Joint Commission on Accreditation of Healthcare Organizations (JCAHO) considers wrong-site surgeries to be *sentinel events* that require immediate investigation and response. As of March 2000, JCAHO has reported wrong-site surgery to be the fourth most commonly reported sentinel event, following patient suicide, medication error, and operative or postoperative complications. It seems inconceivable that this type of error, which carries potentially devastating consequences, could become a common occurrence in organizations comprised of so many highly trained practitioners. While human fallibility, as always, plays a fundamental role, its interplay with contextual factors and existing barriers suggests that these errors are more complex than they appear.

A common factor in wrong-site surgery is the involvement of multiple surgeons on a case. Each of the various physicians has a relatively narrow focus of attention (e.g., the cardiologist is focused on whether the heart can withstand surgery), which decreases the likelihood that the patient will be surrounded by health care providers who are knowledgeable about the case and thus limits the benefits of human redundancy. Another factor in wrong-site surgery is the need to perform multiple procedures during a single trip to the operating room. This factor provides the necessary distractions for a slip. The likelihood that a distraction could result in an unintended action is increased to the extent that the surgeon has "frequently" or "recently" performed surgeries at the unintended site or when patient care is transferred to another surgeon. Fatigue, sleep deprivation, and unusual patient characteristics such as massive obesity (which could alter the positioning of the patient) are also capable of promoting unintended actions by disrupting the surgeon's focused attention.

Presurgical procedures and problems with the way that team members communicate during an operation can also contribute to the occurrence of wrong-site errors. Ideally, an entire team should be required to verify that the correct patient and the correct limb have been prepared for surgery. However, when the surgical team fails to review the patient record or image data in the time period immediately prior to the surgery, memory concerning the correct surgical site can become flawed. Incomplete or inaccurate communication among surgical team members can also occur when some team members are excluded from participating in the site verification process, team members exchange roles during the day of surgery, or when the entire team depends exclusively on the surgeon to identify the surgical site (the latter often occurs in work cultures that accept the surgeon's decision as final). Many of these communication problems become magnified under time constraints stemming from pressure from hospital administrators to speed things up.

A tactic that has recently received considerable attention is marking the operative site and involving the patient in the process. However, even this seemingly straightforward policy can be problematic. If surgeons were to employ their own marking techniques, such as "No" on the wrong limb or "Yes" on the proper site, confusion may occur to the point of increasing the likelihood of wrong-site surgery. Standardization is thus critical, and the recommended procedure is for the surgeon to initialize the operative site. This barrier alone, however, is insufficient. For example, if the marked site is draped out of the surgeon's field of view and the surgeon does not recall whether the site was or was not marked, the possibility for error still exists. Thus, a verification checklist should also be in place that includes all documents referencing the intended procedure and site, informed consent, and direct observation of the marked operative site. Strict reliance on x-rays or the patient's chart can prove inadequate in cases where the data are incorrect or associated with the wrong patient, and patient involvement is not always possible, depending on the patient's condition.

Violation of these barriers is not uncommon. Some surgeons see signing as a waste of time and a practice that could contaminate the operative site. They are insistent that wrong-site surgery errors would not happen to them and that the focus of the medical profession should be on ridding itself of incompetent surgeons rather than instituting wide-reaching programs (Prager, 1998). This attitude, however, is not surprising in a profession that has relied largely on people avoiding mistakes rather than creating systems to minimize them. It also reflects a very traditional perspective to human error whereby the responsibility or blame for errors is placed solely on those who committed them and suggests that a culture shift among surgeons may be needed. To utilize the surgeon's time more efficiently, for hospitalized patients implementation might involve the operating surgeon initializing the intended operative site at the time consent is obtained, thus requiring that the physician be present during consent. The JCAHO has constructed a universal protocol for eliminating wrong-site surgery which ensures that the surgical site is marked while the patient is conscious and that there is a final pause and verification among all surgical team members to ensure that everyone is in agreement with the procedure. This protocol became effective in July 2004 for all JCAHO-accredited hospitals.

4 ERROR TAXONOMIES AND PREDICTING HUMAN ERROR

Many areas of scientific investigation use classification systems or taxonomies as a way of organizing knowledge about a subject matter. In the case of human error, the taxonomies that have been proposed have theoretical as well as practical value. The taxonomies that emphasize observable behaviors are primarily of practical value. They can be used retrospectively to gather data on trends that point to weaknesses in design, training, and operations, as well as prospectively, in conjunction with detailed analyses of tasks and situational contexts, to predict possible errors and to suggest countermeasures for detecting, minimizing, or eliminating these errors. Human error taxonomies can also be directed at specific tasks or operations. For example, a taxonomy could be developed for the purpose of characterizing all the various observable ways that a particular task can be performed incorrectly, analogous to the use of failure mode and effects analysis (Kumamoto and Henley, 1996) to identify a component's failure modes and their corresponding causes and effects. In the health care industry, the diversity of medical procedures and the variety of circumstances under which these procedures are performed may, in fact, call for highly specific error taxonomies.

For more cognitively complex tasks, it may be possible to classify errors according to stages of information processing (Figure 2), thereby differentiating errors related to perception from errors related to failures in working memory. However, many of these errors of cognition can only be inferred from assumptions concerning the human's goals and observed behaviors, and to some extent from contextual factors. The characterization of performance as skill-, rule-, or knowledge-based (Section 3.2.5) has proven particularly useful in thinking about the ways in which information-processing failures can arise, in light of the distinctions in information-processing activities that are presumed to occur at each of these levels. Generally, taxonomies that focus on the cognitive or causal end of the error spectrum have the ability to propose types of errors that might occur under various circumstances and thus can shape or augment our understanding of human limitations in information processing.

A very simple error taxonomy that bears a long history (Sanders and McCormick, 1993) differentiates errors of omission (forgetting to do something) from errors of commission (doing something incorrectly). Errors of commission are often further categorized into errors related to sequence, timing, substitution, and actions not included in a person's current plans (Hollnagel, 1993). *Sequence errors* include actions that are repeated (which may result in restarting a process) or are reversed (which may result in jumping ahead in a sequence). *Timing errors* refer to actions that do not occur when they are required; thus they may occur prematurely or after some delay. *Substitution errors* refer to single actions or sets of actions that are performed in place of the expected action or action set. Errors involving the inclusion of additional actions are referred to as *intrusions* when they are capable of disrupting the planned sequence of actions. Disruptions can lead to *capture* by the sequence, *branching* to an incorrect sequence, or *overshooting* the action sequence beyond the satisfaction of its objective.

Figure 5 and Tables 2 to 4 illustrate several other error taxonomies. The flowchart in Figure 5 classifies different types of human errors that can occur under skill-, rule-, and knowledge-based levels of performance. This flowchart seeks to answer questions concerning how an error occurred. Similar flowcharts are provided by the author to address the more preliminary issue in the causal chain (i.e., why an error occurred) as well as the external manifestation of the error (i.e., what type of error occurred). Reason's (1990) taxonomy (Table 2) also exploits the distinctions among skill-, rule-, and knowledge-based levels of performance, but draws attention to how error modes related to skill-based slips and lapses differ from error modes related to rule- and knowledge-based mistakes. The taxonomies presented in Tables 3 and 4 demonstrate various schemes for classifying errors based on stages of information processing.

In addition to their usefulness for analyzing accidents for root causes (Section 10.2), error taxonomies that emphasize cognitive or causal factors have predictive valuѣ as well. Predicting human error, however, is a difficult matter. It may indeed be possible to construct highly controlled experimental tasks that "trap" people into particular types of skill-based slips and lapses and some forms of rule-based mistakes. However, the multidimensional complexity surrounding actual work situations and the uncertainty associated with the human's goals, intentions, and emotional and attentional states introduce many layers of guesswork into the process of establishing reliable mappings between human fallibility and context. In 1991, Senders and Moray stated: "To understand and predict errors . . . usually requires a detailed task analysis" (p. 60). Nothing has changed since to diminish the validity of this assertion. In fact, the current emphasis on *cognitive task analysis* (CTA) techniques and our greater understanding of mechanisms underlying human error have probably made the process of predicting human error more laborious than ever, as it should be. Expectations of shortcuts are unreasonable; error prediction by its very nature should be a tedious process and will often be influenced by the choice of taxonomy.

Task analysis (TA), which is fundamental to error prediction, describes the human's involvement with a system in terms of task requirements, actions, and cognitive processes (Chapter 14). It can be used to provide a broad overview of task requirements (that are often useful during the preliminary stages of product design) or a highly detailed description of activities. These descriptions could include time constraints and activity time lines; sequential dependencies among activities; alternative plans for performing an operation; contingencies that may arise during the course of activities and options for handling these contingencies; the feedback available at each step of the process; characterizations of information flow between different subsystems; and descriptions of displays, controls, training, and interactions with other people. Tabular formats are often used to illustrate the various relationships between these factors and task activities. Many different TA methods exist (Kirwan and Ainsworth, 1992; Luczak, 1997; Shepherd, 2000) and identifying an appropriate method for a particular problem or work domain can be critical.

In CTA, the interest is in determining how the human conceptualizes tasks, recognizes critical information and patterns of cues, assesses situations, makes discriminations, and uses strategies for solving problems, forming judgments, and making decisions. Successful application of CTA for enhancing system performance will depend on a concurrent understanding of the cognitive processes underlying human performance in the work domain and the constraints on cognitive processing the work domain imposes (Vicente, 1999). In developing new systems, meeting this objective may require multiple, coordinated approaches. As Potter et al. (1998) have noted: "No one approach can capture the richness required for a comprehensive, insightful CTA" (p. 395).

As with TA, many different CTA techniques are presently available (Hollnagel, 2003). TA and CTA, however, should not be viewed as mutually exclusive enterprises—in fact, the case could be made that TA methods that incorporate CTA represent "good" task analyses. As anticipating the precise time and

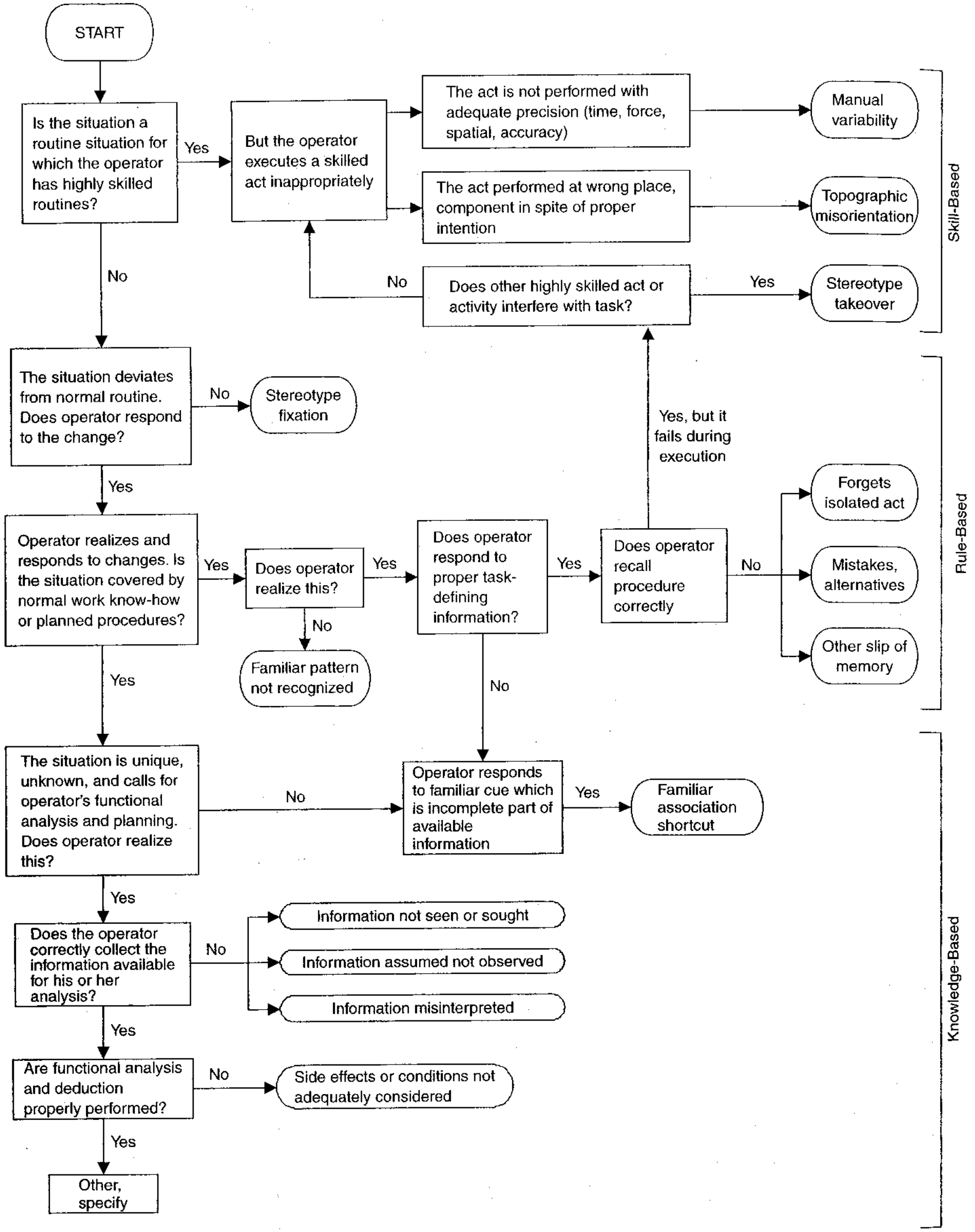

START
Is the situation a routine situation for which the operator has highly skilled routines?
Yes
But the operator executes a skilled act inappropriately
The act is not performed with adequate precision (time, force, spatial, accuracy)
Manual variability
The act performed at wrong place, component in spite of proper intention
Topographic misorientation
No
Does other highly skilled act or activity interfere with task?
Yes
Stereotype takeover
Skill-Based
No
The situation deviates from normal routine. Does operator respond to the change?
No
Stereotype fixation
Yes, but it fails during execution
Yes
Operator realizes and responds to changes. Is the situation covered by normal work know-how or planned procedures?
Yes
Does operator realize this?
Yes
Does operator respond to proper task-defining information?
Yes
Does operator recall procedure correctly
No
Forgets isolated act
Mistakes, alternatives
Other slip of memory
Rule-Based
No
Familiar pattern not recognized
No
Yes
The situation is unique, unknown, and calls for operator's functional analysis and planning. Does operator realize this?
No
Operator responds to familiar cue which is incomplete part of available information
Yes
Familiar association shortcut
Yes
Does the operator correctly collect the information available for his or her analysis?
No
Information not seen or sought
Information assumed not observed
Information misinterpreted
Yes
Are functional analysis and deduction properly performed?
No
Side effects or conditions not adequately considered
Yes
Other, specify
Knowledge-Based

Figure 5 Decision flow diagram for analyzing an event into one of 13 types of human error. (From Rasmussen, 1982; copyright 1982, with permission from Elsevier.)

Table 2 Human Error Modes Associated with Rasmussen's SRK Framework

Skill-Based Performance	
Inattention	Overattention
Double-capture slips Omissions following interruptions Reduced intentionality Perceptual confusions Interference errors	Omissions Repetitions Reversals
Rule-Based Performance	
Misapplication of Good Rules	Application of Bad Rules
First exceptions Countersigns and nonsigns Informational overload Rule strength General rules Redundancy Rigidity	Encoding deficiencies Action deficiencies Wrong rules Inelegant rules Inadvisable rules
Knowledge-Based Performance	
Selectivity Workspace limitations Out of sight, out of mind Confirmation bias Overconfidence Biased reviewing Illusory correlation Halo effects Problems with causality	Problems with complexity Problems with delayed feedback Insufficient consideration of processes in time Difficulties with exponential developments Thinking in causal series and not causal nets Thematic vagabonding Encysting

Source: Reason (1990).

mode of error is generally unrealistic, the use of TA techniques should be directed at uncovering the possibility for errors and prioritizing these possibilities. Given what we can surmise about human fallibility, the contexts within which human activities occur, and the barriers that may be in place, the relevant questions are then as follows: What kinds of actions by people are possible or even reasonable that would, by one's definition, constitute errors? What are the possible consequences of these errors? What kinds of barriers do these errors and their consequences call for? Depending on whether the analysis is to be applied to a product or process that is still in the conceptual stages, to a newly implemented process, or to an existing process, broad applications of TA techniques that may include mock-ups, walkthroughs, simulations, interviews, and direct observations are needed to identify the relevant contextual elements. In-depth task analyses that incorporate CTA techniques could then provide the details necessary for evaluating the various possibilities for interplay between context and human fallibility (Sharit, 1998).

Table 3 External Error Modes Classified According to Stages of Human Information Processing

1. Activation/detection
   - 1.1 Fails to detect signal/cue
   - 1.2 Incomplete/partial detection
   - 1.3 Ignore signal
   - 1.4 Signal absent
   - 1.5 Fails to detect deterioration of situation
2. Observation/data collection
   - 2.1 Insufficient information gathered
   - 2.2 Confusing information gathered
   - 2.3 Monitoring/observation omitted
3. Identification of system state
   - 3.1 Plant-state-identification failure
   - 3.2 Incomplete-state identification
   - 3.3 Incorrect-state identification
4. Interpretation
   - 4.1 Incorrect interpretation
   - 4.2 Incomplete interpretation
   - 4.3 Problem solving (other)
5. Evaluation
   - 5.1 Judgment error
   - 5.2 Problem-solving error (evaluation)
   - 5.3 Fails to define criteria
   - 5.4 Fails to carry out evaluation
6. Goal selection and task definition
   - 6.1 Fails to define goal/task
   - 6.2 Defines incomplete goal/task
   - 6.3 Defines incorrect or inappropriate goal/task
7. Procedure selection
   - 7.1 Selects wrong procedure
   - 7.2 Procedure inadequately formulated/shortcut invoked
   - 7.3 Procedure contains rule violation
   - 7.4 Fails to select or identify procedure
8. Procedure execution
   - 8.1 Too early/late
   - 8.2 Too much/little
   - 8.3 Wrong sequence
   - 8.4 Repeated action
   - 8.5 Substitution/intrusion error
   - 8.6 Orientation/misalignment error
   - 8.7 Right action on wrong object
   - 8.8 Wrong action on right object
   - 8.9 Check omitted
   - 8.10 Check fails/wrong check
   - 8.11 Check mistimed
   - 8.12 Communication error
   - 8.13 Act performed wrongly
   - 8.14 Part of act performed
   - 8.15 Forgets isolated act at end of task
   - 8.16 Accidental timing with other event/circumstance
   - 8.17 Latent error prevents execution
   - 8.18 Action omitted
   - 8.19 Information not obtained/transmitted
   - 8.20 Wrong information obtained/transmitted
   - 8.21 Other

Source: Kirwan (1994).

Even when applied at relatively superficial levels, TA techniques are well suited for identifying mismatches between demands imposed by the work

Table 4 Human Error Classification Scheme

1. Observation of system state
    - Improper rechecking of correct readings
    - Erroneous interpretation of correct readings
    - Incorrect readings of appropriate state variables
    - Failure to observe sufficient number of variables
    - Observation of inappropriate state variables
    - Failure to observe any state variables
2. Choice of hypothesis
    - Hypotheses could not cause the values of the state variables observed
    - Much more likely causes should be considered first
    - Very costly place to start
    - Hypothesis does not functionally relate to the variables observed
3. Testing of hypothesis
    - Stopped before reaching a conclusion
    - Reached wrong conclusion
    - Considered and discarded correct conclusion
    - Hypothesis not tested
4. Choice of goal
    - Insufficient specification of goal
    - Choice of counterproductive or nonproductive goal
    - Goal not chosen
5. Choice of procedure
    - Choice would not fully achieve goal
    - Choice would achieve incorrect goal
    - Choice unnecessary for achieving goal
    - Procedure not chosen
6. Execution of procedure
    - Required stop omitted
    - Unnecessary repetition of required step
    - Unnecessary step added
    - Steps executed in wrong order
    - Step executed too early or too late
    - Control in wrong position or range
    - Stopped before procedure complete
    - Unrelated inappropriate step executed

Source: Rouse and Rouse (1983).

context and the human's capabilities for meeting these demands. Although hypothesizing specific error forms will become more difficult at this level of analysis, windows of opportunity for error still can be readily exposed that, in and of themselves, can suggest countermeasures capable of reducing risk potential. For example, these analyses may determine that there is insufficient time to input information accurately into a computer-based documentation system, that the design of displays is likely to evoke control responses that are contraindicated, or that sources of information on which high-risk decisions are based contain incomplete or ambiguous information. This coarser approach to predicting errors or error-inducing conditions that derives from analyzing demand-capability mismatches can also highlight contextual and cognitive considerations that can form the basis for a more focused application of TA and CTA techniques.

Table 5 depicts a portion of a type of TA known as a *hierarchical task analysis* (HTA) that was developed

Table 5 Part of a Hierarchical Task Analysis Associated with Filling a Chlorine Tanker

0. Fill tanker with chlorine.
Plan: Do tasks 1 to 5 in order.
1. Park tanker and check documents (not analyzed).
2. Prepare tanker for filling.
Plan: Do 2.1 or 2.2 in any order, then do 2.3 to 2.5 in order.
2.1 Verify tanker is empty.
Plan: Do in order:
2.1.1 Open test valve.
2.1.2 Test for Cl_2.
2.1.3 Close test valve.
2.2 Check weight of tanker.
2.3 Enter tanker target weight.
2.4 Prepare fill line.
Plan: Do in order:
2.4.1 Vent and purge line.
2.4.2 Ensure main Cl_2 valve is closed.
2.5 Connect main Cl_2 fill line.
3. Initiate and monitor tanker filling operation.
Plan: Do in order:
3.1 Initiate filling operation.
Plan: Do in order:
3.1.1 Open supply line valves.
3.1.2 Ensure tanker is filling with chlorine.
3.2 Monitor tanker filling operation.
Plan: Do 3.2.1, do 3.2.2 every 20 minutes;
on initial weight alarm, do 3.2.3 and 3.2.4;
on final weight alarm, do 3.2.5 and 3.2.6.
3.2.1 Remain within earshot while tanker is filling.
3.2.2 Check road tanker.
3.2.3 Attend tanker during last filling of 2 or 3 tons.
3.2.4 Cancel initial weight alarm and remain at controls.
3.2.5 Cancel final weight alarm.
3.2.6 Close supply valve A when target weight is reached.
4. Terminate filling and release tanker.
4.1 Stop filling operation.
Plan: Do in order:
4.1.1 Close supply valve B.
4.1.2 Clear lines.
4.1.3 Close tanker valve.
4.2 Disconnect tanker.
Plan: Repeat 4.2.1 five times, then do 4.2.2 to 4.2.4 in order.
4.2.1 Vent and purge lines.
4.2.2 Remove instrument air from valves.
4.2.3 Secure blocking device on valves.
4.2.4 Break tanker connections.
4.3 Store hoses.
4.4 Secure tanker.
Plan: Do in order:
4.4.1 Check valves for leakage.
4.4.2 Secure log-in nuts.
4.4.3 Close and secure dome.
4.5 Secure panel (not analyzed).
5. Document and report (not analyzed).

Source: CCPS (1994). Copyright 1994 by the American Institute of Chemical Engineers, and reproduced by permission of AIChE.

for analyzing the task of filling a storage tank with chlorine from a tank truck. The primary purpose of this TA was to identify potential human errors that could contribute to a major flammable release resulting either from a spill during unloading of the truck or from a tank rupture. Table 6 illustrates the use of this HTA for predicting external error modes. The error taxonomy shown in Table 3 can easily be adapted for predicting the types of errors listed in Table 6. This taxonomy can also be linked to more underlying psychological mechanisms, allowing errors with identical or similar external manifestations to be distinguished and thus adding considerable depth to the understanding of potential errors predicted from the TA. As discussed in Section 5.4, this ability not only results in more accurate quantification of error data but also provides the basis for more effective error-reduction strategies. An example of such a scheme is the *human error identification in systems technique* (HEIST), which classifies external error modes according to the eight stages of human information processing listed in Table 3. The first column in a HEIST table consists of a code whose initial letter(s) refers to one of these eight stages. The next letter in the code refers to one of six general PSFs: time (T), interface (I), training/experience/familiarity (E), procedures (P), task organization (O), and task complexity (C). The external error modes are then linked to underlying psychological error mechanisms (PEMs). Many of these mechanisms are consistent with the failure modes that appear in Reason's error taxonomy (Table 2).

Table 7 presents an extract from a HEIST table corresponding to two of the eight stages of human information processing listed in Table 3: activation/detection and observation/data collection. For these two stages of information processing, more detailed explanations of the PEMs listed in the HEIST table may be found in Table 8. A complete HEIST table and the corresponding listing of PEMs can be found in Kirwan (1994).

On a final note, task analysts contending with complex systems will often need to consider various properties of the wider system or subsystem in which human activities take place. As Shepherd (2000) has stated: "Any task analysis method which purports to serve practical ends needs to be carried out beneath a general umbrella of systems thinking" (p. 11). There are a variety of ways in which systems can be characterized or decomposed (Sharit, 1997), and for any particular system these various descriptions could lead to the consideration of different activities for analysis as well as different strategies for performing these analyses.

5 QUANTIFYING HUMAN ERROR

5.1 Historical Antecedents

Quantifying human error presumes that a probability can be attached to its occurrence. Is this a realistic endeavor? An objective assignment of probabilities to events requires that *human error probability* (HEP) be defined as a ratio of the number of observed occurrences of the error to the number of opportunities for that error to occur. Based on this definition, it can be argued that with the exception of routine skill-based activities, estimates of HEPs are not easily attainable or likely to be accurate. Assuming that most organizations would be more interested in gauging the possibility for human error and understanding its causality and consequences, the more compelling question is: Why do we need a quantitative estimate?

The catalyst behind quantification of human error was the mandate for industries involved in high-hazard operations to perform *probabilistic risk analyses* (PRAs). Most industries that carry out such assessments, such as the chemical processing and nuclear power industries, are concerned about hazards arising from interactions among various system events, including hardware and software failures, environmental anomalies, and human errors that are capable of producing injuries, fatalities, disruptions to production, and plant and environmental damage. The two primary hazard analysis techniques that have become associated with PRAs are *fault tree* (FT) analysis and *event tree* (ET) analysis. The starting point for each of these methods is an undesirable event. Other hazard analysis techniques (CCPS, 1992) or methods based on expert opinion are often used to identify these events. FTs utilize Boolean logic models to depict the relationships among hardware, human, and environmental events that can lead to the undesirable *top event*. When FTs are used as a quantitative method, *basic events* (for which no further analysis of the cause is carried out) are assigned probabilities or occurrence rates, which are then propagated into a probability or rate measure associated with the top event (Dhillon and Singh, 1981). The contributions of each of the singular events to the top event can also be computed, making this technique very suitable for cost–benefit analyses that can be used as a basis for specifying design interventions. As a qualitative analysis tool, FTs can identify the various combinations of events (or cut sets) that could lead to the top event; for many applications this information is sufficiently revealing for satisfying safety objectives.

Whereas a FT represents a deductive, top-down decomposition of an undesirable event (such as a loss in electrical power), an ET corresponds to an inductive analysis that determines how this undesirable event can propagate. These trees are thus capable of depicting the various sequences of events that can become triggered by the initiating event, as well as the risks associated with each of these sequences. Figure 6 illustrates a simple ET consisting of two operator actions and two safety systems. When ETs are constructed to address only sequences of human actions in response to the initiating event, the ET is sometimes referred to as an *operator action event tree* (OAET). In OAETs, each branch of the tree represents either a success or an HEP associated with the required actions specified along the column headings. These trees can easily accommodate paths signifying recovery from previous errors. In many

Table 6 Human Errors and Error Reduction Recommendations for the HTA in Table 5[a]

				Error Reduction Recommendations		
Error Type	Error Description	Recovery	Consequences and Comments	Procedures	Training	Equipment
Wrong information obtained	Wrong weight is entered.	On check	Alarm does not sound before tanker overfills.	Validate target weight independently.	Ensure that operator double-checks data entered; record values in checklist.	Provide automatic setting of weight alarms from unladen weight; install computerized logging system and built-in checks on tanker reg. no. and unladen weight linked to warning system; display differences.
Check omitted	Tanker is not monitored while filling.	On initial weight alarm	Alarm will alert the operator if set correctly. Equipment fault (e.g., leaks not detected early and remedial action delayed).	Provide secondary task involving other personnel; supervisor checks operation periodically.	Stress importance of regular checks for safety.	Provide automatic log-in procedure.
	Operator fails to attend.	On step 3.2.5	If alarm is not detected within 10 minutes, tanker will overfill.	Ensure work schedule allows operator to do this without pressure.	Illustrate consequences of not attending.	Repeat alarm in secondary area; provide automatic interlock to terminate loading if alarm is not acknowledged; provide visual indication of alarm.
	Final weight alarm is taken as initial weight alarm.	No recovery	Tanker overfills.	Note differences between the sound of the two alarms in checklist.	Alert operators during training about differences in sounds of alarms.	Use completely different tones for initial and final weight alarms.
	Tanker valve is not closed.	On step 4.2.1	Failure to close tanker valve would result in pressure not being detected during the pressure check in step 4.2.1.	Perform independent check on action; use checklist.	Ensure that operator is aware of consequences of failure.	Valve position indicator would reduce probability of error.
Operation omitted, operation incomplete	Lines are not fully purged.	On step 4.2.4	Failure of operator to detect pressure in lines could lead to leak when tanker connections are broken	Specify a procedure to indicate how to check if fully purged.	Ensure that training covers symptoms of pressure in line.	Line pressure indicator used at controls, interlock device on line pressure.
Operation omitted	Locking nuts are left unsecured.	None	Failure to secure locking nuts could result in leakage during transportation.	Use checklist.	Stress safety implication of training.	Locking nuts gives tactile feedback when secure.

Source: Adapted from CCPS (1994). Copyright 1994 by the American Institute of Chemical Engineers, and reproduced by permission of AIChE.

[a] Possible errors derive from Table 3.

Table 7 Extract from a HEIST Table

Code	Error-Identifier Prompt	External Error Mode	System Cause/Psychological Error-Mechanism	Error-Reduction Guidelines
AT1	Does the signal occur at the appropriate time? Could it be delayed?	Action omitted or performed either too early or too late	Signal timing deficiency, failure of prospective memory	Alter system configuration to present signal appropriately; generate hard copy to aid prospective memory; repeat signal until action has occurred.
AI1	Could the signal source fail?	Action omitted or performed too late	Signal failure	Use diverse/redundant signal sources; use a higher-reliability signal system; give training and ensure that procedures incorporate investigation checks on "no signal."
AI2	Can the signal be perceived as unreliable?	Action omitted	Signal ignored	Use diverse signal sources; ensure higher signal reliability; retrain if signal is more reliable than it is perceived to be.
AI3	Is the signal a strong one, and is it in a prominent location? Could the signal be confused with another?	Action omitted, or performed too late, or wrong act performed	Signal-detection failure	Prioritize signals; place signals in primary (and unobscured) location; use diverse signals; use multiple-signal coding; give training in signal priorities; make procedures cross-reference the relevant signals; increase signal intensity.
AI4	Does the signal rely on oral communication?	Action omitted or performed too late	Communication failure, lapse of memory	Provide physical backup/substitute signal; build required communications requirements into procedures.
AE1	Is the signal very rare?	Action omitted or performed too late	Signal ignored (false alarm), stereotype fixation	Give training for low-frequency events; ensure diversity of signals; prioritize signals into a hierarchy of several levels.
AE2	Does the operator understand the significance of the signal?	Action omitted or performed too late	Inadequate mental model	Training and procedures should be amended to ensure that significance is understood.
AP1	Are procedures clear about action following the signal or the previous step, or when to start the task?	Action omitted or performed either too early or too late	Incorrect mental model	Procedures must be rendered accurate, or at least made more precise; give training if judgment is required on when to act.
AO1	Does activation rely on prospective memory (i.e., remembering to do something at a future time, with no specific cue or signal at that later time)?	Action omitted or performed either too late or too early	Prospective memory failure	Proceduralize task, noting calling conditions, timings of actions, etc.; utilize an interlock system preventing task from occurring at undesirable times; provide a later cue; emphasize this aspect during training.
AO2	Will the operator have other duties to perform concurrently? Are there likely to be distractions? Could the operator become incapacitated?	Action omitted or performed too late	Lapse of memory, memory failure, signal-detection failure	Training should prioritize signal importance; improve task organization for crew; use memory aids; use a recurring signal; consider automation; utilize flexible crewing.
AO3	Will the operator have a very high or low workload?	Action omitted or performed either too late or too early	Lapse of memory, other memory failure, signal-detection failure	Improve task and crew organization; use a recurring signal; consider automation; utilize flexible crewing; enhance signal salience.

(continued overleaf)

Table 7 *(continued)*

Code	Error-Identifier Prompt	External Error Mode	System Cause/Psychological Error-Mechanism	Error-Reduction Guidelines
AO4	Will it be clear who must respond?	Action omitted or performed too late	Crew-coordination failure	Emphasize task responsibility in training and task allocation among crew members; utilize team training.
AC1	Is the signal highly complex?	Action omitted, or wrong act performed either too late or too early	Cognitive overload, inadequate mental model	Simplify signal; automate system response; give adequate training in the nature of the signal; provide online, automated, diagnostic support; develop procedures that allow rapid analysis of the signal (e.g., use of flowcharts).
AC2	Is the signal in conflict with the current diagnostic mindset?	Action omitted or wrong act performed	Confirmation bias, signal ignored	Procedures should emphasize disconfirming as well as confirmatory signals; utilize a shift technical advisory in the shift structure; carry out problem-solving training and team training; utilize diverse signals; implement automation.
AC3	Could the signal be seen as part of a different signal set? Or is, in fact, the signal part of a series of signals to which the operator needs to respond?	Action performed too early or wrong act performed	Familiar-association shortcut/ stereotype takeover	Training and procedures could involve display of signals embedded within mimics or other representations showing their true contexts or range of possible contexts; use fault-symptom matrix aids; etc.
OT1	Could the information or check occur at the wrong time?	Failure to act, or action performed either too late or too early, or wrong act performed	Inadequate mental model/ inexperience/ crew coordination failure	Procedure and training should specify the priority and timing of checks; present key information centrally; utilize trend displays and predictor displays if possible; implement team training.
OI1	Could important information be missing due to instrument failure?	Action omitted or performed either too late or too early, or wrong act performed	Signal failure	Use diverse signal sources; maintain backup power supplies for signals; have periodic manual checks; procedures should specify action to be taken in event of signal failure; engineer automatic protection/action; use a higher-reliability system.
OI2	Could information sources be erroneous?	Action omitted or performed either too late or too early, or wrong act performed	Erroneous signal	Use diverse signal sources; procedures should specify cross-checking; design system-self-integrity monitoring; use higher-reliability signals.
OI3	Could the operator select an incorrect but similar information source?	Action omitted or performed either too late or too early, or wrong act performed	Mistakes alternatives, spatial misorientation, topographic misorientation	Ensure unique coding of displays, cross-referenced in procedures; enhance discriminability via coding; improve training.
OI4	Is an information source accessed only via oral communication?	Action omitted or performed either too late or too early, or wrong act performed	Communication failure	Use diverse signals from hardwired or softwired displays; ensure backup human corroboration; design communication protocols.
OI5	Are any information sources ambiguous?	Action omitted or performed either too late or too early, or wrong act performed	Misinterpretation, mistakes alternatives	Use task-based displays; design symptom-based diagnostic aids; utilize diverse information sources; ensure clarity of information displayed; utilize alarm conditioning.

(continued overleaf)

Table 7 ***(continued)***

Code	Error-Identifier Prompt	External Error Mode	System Cause/Psychological Error-Mechanism	Error-Reduction Guidelines
OI6	Is an information source difficult or time-consuming to access?	Action omitted or performed too late, or wrong act performed	Information assumed	Centralize key data; enhance data access; provide training on importance of verification of signals; enhance procedures.
OI7	Is there an abundance of information in the scenario, some of which is irrelevant, or a large part of which is redundant?	Action omitted or performed too late	Information overload	Prioritize information displays (especially alarms); utilize overview mimics (VDU or hardwired); put training and procedural emphasis on data-collection priorities and data management.
OE1	Could the operator focus on key indication(s) related to a potential event while ignoring other information sources?	Action omitted or performed too late, or wrong act performed	Confirmation bias, tunnel vision	Provide training in diagnostic skills; enhance procedural structuring of diagnosis, emphasizing checks on disconfirming evidence; implement a staff-technical-advisor role; present overview mimics of key parameters showing whether system integrity is improving or worsening or adequate.
OE2	Could the operator interrogate too many information sources for too long, so that progress toward stating identification or action is not achieved?	Action omitted or performed too late	Thematic vagabonding, risk-recognition failure, inadequate mental model	Provide training in fault diagnosis; provide team training; put procedural emphasis on required data collection time frames; implement high-level indicators (alarms) of system-integrity deterioration.
OE3	Could the operator fail to realize the need to check a particular source? Is there an adequate cue prompting the operator?	Action omitted or performed either too late or too early, or wrong act performed	Need for information not prompted, prospective memory failure	Provide procedural guidance on checks required, training, use of memory aids, use of attention-gaining devices (false alarms, central displays, and messages)
OE4	Could the operator terminate the data collection/observation early?	Action omitted or performed either too early or too late, or wrong act performed	Overconfidence, inadequate mental model, incorrect mental model, familiar-association shortcut	Provide training in diagnostic procedures and verification; provide procedural specification of required checks, etc.; implement a shift-technical-advisor role.
OE5	Could the operator fail to recognize that special circumstances apply?	Action omitted or performed either too late or too early, or wrong act performed	Failure to consider special circumstances, slip of memory, inadequate mental model	Ensure training for, as well as procedural noting of, special circumstance; STA; give local warnings in the interface displays/controls.
OP1	Could the operator fail to follow the procedures entirely?	Action omitted or wrong act performed	Rule violation, risk-recognition failure, production–safety conflict, safety-culture deficiency	Provide training in use of procedures; involve operator in development and verification of procedures.

(continued overleaf)

Table 7 *(continued)*

Code	Error-Identifier Prompt	External Error Mode	System Cause/Psychological Error-Mechanism	Error-Reduction Guidelines
OP2	Could the operator forget one or more items in the procedures?	Action omitted or performed either too early or too late, or wrong act performed	Forget isolated act, slip of memory, place-losing error	Ensure an ergonomic procedure design; utilize tick-off sheets, place keeping aids, etc.; provide team training to emphasize checking by other team member(s).
OO1 (AO2)	Will the operator have other duties to perform concurrently? Are there likely to be distractions? Could the operator become incapacitated?	Action omitted or performed too late	Lapse of memory, memory failure, signal-detection failure	Training should prioritize signal importances; develop better task organization for crew; use memory aids; use a recurring signal; consider automation; use flexible crewing.
OO2 (AO3)	Will the operator have a very high or low workload?	Action omitted or performed either too late or too early	Lapse of memory, other memory failure, signal-detection failure	Establish better task and crew organization; utilize a recurring signal; consider automation; use flexible crewing; enhance signal salience.
OO3 (AO4)	Will it be clear who must respond?	Action omitted or performed too late	Crew-coordination failure	Improve training and task allocation among crew; provide team training.
OO4	Could information collected fail to be transmitted effectively across shift-handover boundaries?	Failure to act, or wrong action performed, or action performed either too late or too early, or an error of quality (too little or too much)	Crew-coordination failure	Develop robust shift-handover procedures; training; provide team training across shift boundaries; develop robust and auditable data-recording systems (logs).
OC1	Does the scenario involve multiple events, thus causing a high level of complexity or a high workload?	Failure to act, or wrong action performed, or action performed either too early or too late	Cognitive overload	Provide emergency-response training; design crash-shutdown facilities; use flexible crewing strategies; implement shift-technical-advisor role; develop emergency operating procedures able to deal with multiple transients; engineer automatic information recording (trends, logs, printouts); generate decision/diagnostic support facilities.

Source: Adapted from Kirwan (1994).

PRA applications, FTs and ETs are combined—each major column of the ET can represent a top event whose failure probability can be computed through the evaluation of a corresponding FT model (Figure 7).

Quantitative solutions to FTs or ETs that address only machine or material components are ultimately dictated by well-documented mathematical methods for computing component reliability, either in terms of the probability that the component or subsystem functions normally each time it is used or in terms of the probability that the component will not fail during some prescribed time of use (Kapur and Lamberson, 1977). The realization that human interaction with other system components, including other humans, may have a marked effect on the outcomes of PRAs required developing methods for assessing human reliability, thus establishing the field of *human reliability analysis* (HRA). A variety of methods of HRA are currently available that range from relatively quick assessment procedures to those that involve detailed analyses (Kirwan, 1994). Almost all these methods rely on the idea of PSFs, discussed earlier (Section 3.3); the methods differ, however, in how PSFs are used to generate HEPs for various activities. To illustrate the different approaches to deriving HEPs that these methods can take, two

Table 8 Psychological Error Mechanisms for Two of the Stages of Information Processing Presented in Table 7

Activation/Detection

1. *Vigilance failure:* lapse of attention. Ergonomic design of interface to allow provision of effective attention-gaining measures; supervision and checking; task-organization optimization, so that the operators are not inactive for long periods and are not isolated.
2. *Cognitive/stimulus overload:* too many signals present for the operator to cope with. Prioritization of signals (e.g., high-, medium-, and low-level alarms); overview displays; decision-support systems; simplification of signals; flowchart procedures; simulator training; automation.
3. *Stereotype fixation:* operator fails to realize that situation has deviated from norm. Training and procedural emphasis on range of possible symptoms/causes; fault-symptom matrix as a job aid; decision support system; shift technical advisor/supervision.
4. *Signal unreliable:* operator treats signal as false due to its unreliability. Improved signal reliability; diversity of signals; increased level of tolerance on the part of the system, or delay in effects of error, which allows error detection and correction (decreases "coupling"); training in consequences associated with incorrect false-alarm diagnosis.
5. *Signal absent:* signal absent due to a maintenance/calibration failure or a hardware/software error. Provide signal; redundancy/diversity in signaling-design approach; procedures/training to allow operator to recognize when signal is absent.
6. *Signal-discrimination failure:* operator fails to realize that the signal is different. Improved ergonomics in the interface design; enhanced training and procedural support in the area of signal differentiation; supervision checking.

Observation/Data Collection

7. *Attention failure:* lapse of attention.
8. *Multiple signal coding:* enhanced alarm salience; improved task organization with respect to backup crew and rest pauses.
9. *Inaccurate recall:* operator remembers data incorrectly (usually, quantitative data). Nonreliance on memorized data, which would necessitate better interface design – as data are received, they can either be acted on while still present on a display (controls and displays are co-located) or at least be logged onto a "scratch pad"; sufficient displays for presenting all information necessary for a decision/action simultaneously; printer usage; training in nonreliance on memorized data.
10. *Confirmation bias:* operator only selects data that confirm given hypothesis and ignores other disconfirming data sources. Problem-solving training; team training [including training in the need to question decisions, and in the ability of the team leader(s) to take constructive criticism]; shift technical advisor (diverse, highly qualified operator who can "stand back" and consider alternative diagnoses), functional procedures: high-level information displays; simulator training; high-level alarms for system-integrity degradation; automatic protection.
11. *Thematic vagabonding:* operator flits from datum to datum, never actually collating it meaningfully. Problem-solving training; team training; simulator training; functional-procedure specification for decision-timing requirements; high-level alarms for system-integrity degradation.
12. *Encystment:* operator focuses exclusively on only one data source. Problem-solving training; team training [including training in the need to question decisions and in the ability of the team leader(s) to take constructive criticism]; shift technical advisor; functional procedures; high-level information displays; simulator training; high-level alarms for system-integrity degradation.
13. *Stereotype fixation revisited:* need for information is not prompted by either memory or procedures. Emergency procedure enhancements, and emphasis of key symptoms and indicators to be checked; team training; problem-solving training; alarm reprioritization; simulator training.
14. *Crew-functioning problem:* allocation of responsibility or priorities is unclear, with the result that data collection/observation fails.
15. *Cognitive/stimulus overload:* operator too busy, or being bombarded by signals, with the result that effective data collection/observation fails. See item 2.

Source: Kirwan (1994).

well-known techniques that were originally developed for application in the nuclear power industry will be described.

5.2 THERP

The *technique for human error rate prediction*, generally referred to as THERP, is detailed in a work by Swain and Guttmann (1983) sponsored by the U.S. Nuclear Regulatory Commission. Its methodology is driven by decomposition: Human tasks are first decomposed into clearly separable actions or subtasks; HEP estimates are then assigned to each of these actions; and finally, these HEPs are aggregated to derive probabilities of task failure, which reflect human reliability.

THERP is a highly systematic procedure. Its initial steps are directed at establishing which work activities will require emphasis and the time and skill requirements and concerns for human error associated with these activities. Factors related to error detection and the potential for error recovery are also determined. The results of these efforts are represented by a type of event tree referred to as a *probability tree*. Each relevant subtask in a probability

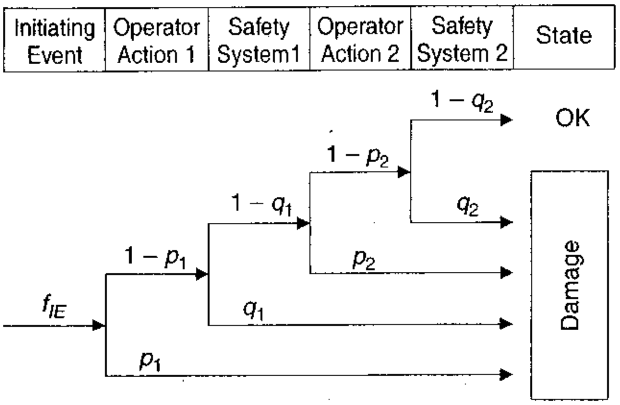


Figure 6 Event tree, where f_{IE} represents the probability of the initiating event, p_1 and p_2 represent human error probabilities associated with two operator actions, and q_1 and q_2 represent system failure probabilities for two safety systems. Typically, the "damage" consequence in the last column is stated more specifically in terms of different accident possibilities. Also, depending on the "level" of the risk analysis, the last column could be extended to reflect the different consequences of the various possible accidents. (From Kumamoto and Henley, 1996; © 2004 IEEE.)

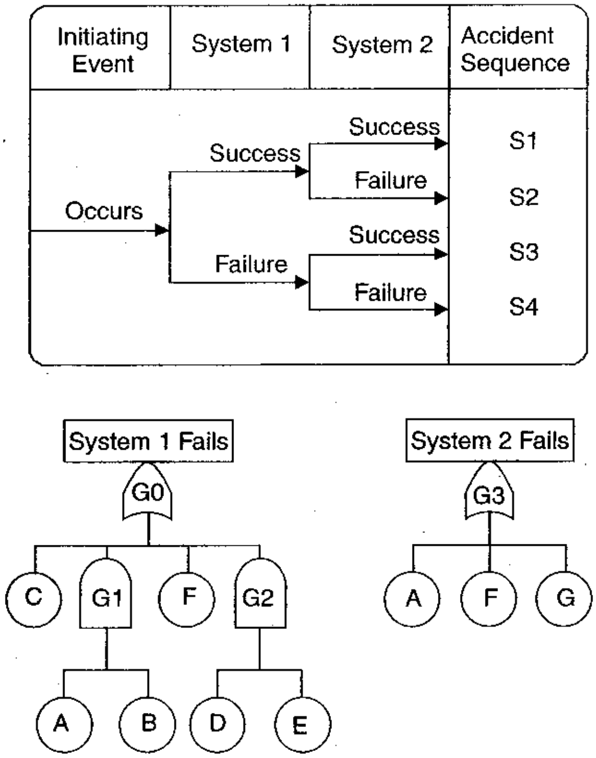


Figure 7 Coupling of event trees and fault trees. The probabilities of failure associated with systems 1 and 2 in the event tree would be derived from the two corresponding fault tees. (From Kumamoto and Henley, 1996; © 2004 IEEE.)

tree is characterized by two limbs, representing either successful or unsuccessful performance (Figure 8).

The next set of steps in THERP constitutes the quantitative assessment stage. First, HEPs are assigned to each of the limbs of the tree corresponding to incorrect performance. These probabilities, referred to as *nominal HEPs*, in theory are presumed to represent medians of lognormal probability distributions. Associated with each nominal HEP are *upper and lower uncertainty bounds* (UCBs), which reflect the variance associated with any given error distribution. The square root of the ratio of the upper to the lower UCB defines the *error factor* (the value selected for this factor will depend on the variability believed to be associated with the probability distribution for that error). Swain and Guttmann (1983) provide values of nominal HEPs and their corresponding error factors for a variety of nuclear power plant tasks. For some tasks the nominal HEPs that are provided refer to *joint HEPs* because it is the performance of a team rather than that of an individual worker that is being evaluated. Generally, the absence of existing hard data from the operations of interest will require that nominal HEPs be derived from other sources, which include (1) expert judgment elicited through techniques such as direct numerical estimation or paired comparisons (Swain and Guttmann, 1983; Kirwan, 1994), (2) simulators (Gertman and Blackman, 1994), and (3) data from jobs similar in psychological content to the operations of interest.

To account for more specific individual-, environmental-, and task-related influences on performance, nominal HEPs are subjected to a series of refinements. First, nominal HEPs are modified based on the influence of PSFs, resulting in *basic HEPs* (BHEPs). In some cases, guidelines are provided in tables indicating the direction and extent of influence of particular PSFs on nominal HEPs; for example, adjustments that are to be made in nominal HEPs due to the influence of the PSF of stress are provided as a function of type of task and worker experience. Next, a nonlinear dependency model is incorporated which considers positive dependencies that exist between adjacent limbs of the tree, resulting in *conditional HEPs* (CHEPs). In a positive dependency model, failure on a subtask increases the probability of failure on the following subtask, and successful performance of a subtask decreases the probability of failure in performing the subsequent task element. Instances of negative dependence can be accounted for but require the discretion of the analyst. In the case of positive dependence, THERP provides equations for modifying BHEPs to CHEPs based on the extent to which the analyst believes dependencies exist.

At this point, success and failure probabilities are computed for the entire task. Various approaches to these computations can be taken. The most straightforward approach is to multiply the individual CHEPs associated with each path on the tree leading to failure, sum these individual failure probabilities to arrive at the probability of failure for the total task, and then assign UCBs to this probability. More complex approaches to these computations take into account the variability associated with the combinations of

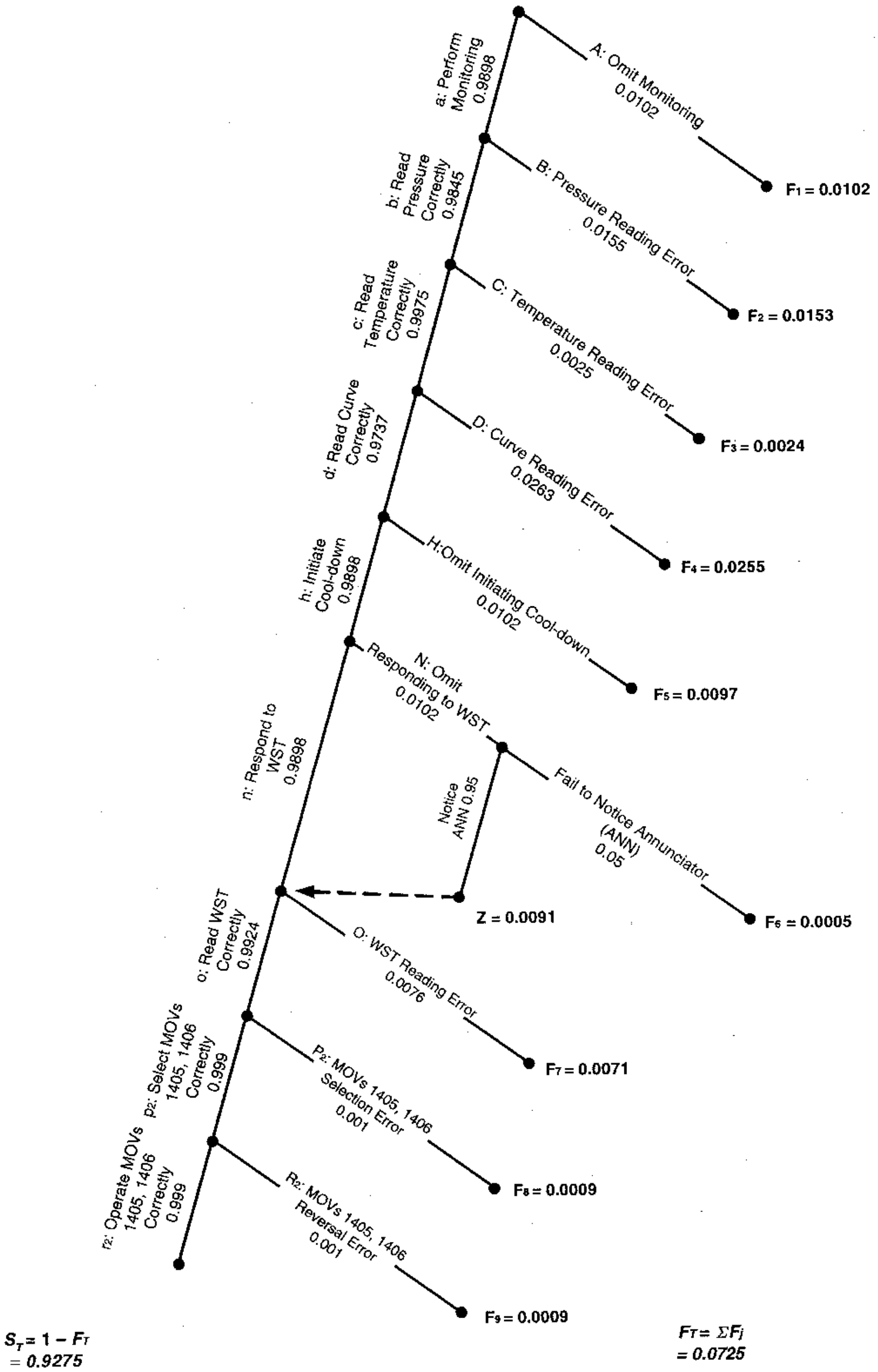

a: Perform Monitoring 0.9898
A: Omit Monitoring 0.0102
F1 = 0.0102
b: Read Pressure Correctly 0.9845
B: Pressure Reading Error 0.0155
F2 = 0.0153
c: Read Temperature Correctly 0.9975
C: Temperature Reading Error 0.0025
F3 = 0.0024
d: Read Curve Correctly 0.9737
D: Curve Reading Error 0.0263
F4 = 0.0255
h: Initiate Cool-down 0.9898
H:Omit Initiating Cool-down 0.0102
F5 = 0.0097
n: Respond to WST 0.9898
N: Omit Responding to WST 0.0102
Notice ANN 0.95
Fail to Notice Annunciator (ANN) 0.05
F6 = 0.0005
Z = 0.0091
o: Read WST Correctly 0.9924
O: WST Reading Error 0.0076
F7 = 0.0071
p2: Select MOVs 1405, 1406 Correctly 0.999
P2: MOVs 1405, 1406 Selection Error 0.001
F8 = 0.0009
r2: Operate MOVs 1405, 1406 Correctly 0.999
R2: MOVs 1405, 1406 Reversal Error 0.001
F9 = 0.0009
ST = 1 – FT = 0.9275
FT = ΣFi = 0.0725

Figure 8 HRA event tree corresponding to a nuclear power control room task that includes one recovery factor. (From Kumamoto and Henley, 1996; © 2004 IEEE.)

events comprising the probability tree (Swain and Guttmann, 1983).

The final steps of THERP consider the ways in which errors can be recovered and the kinds of design interventions that can have the greatest impact on task success probability. Common recovery factors include the presence of annunciators that can alert the operator to the occurrence of an error, co-workers potentially capable of catching or discovering (in time) a fellow worker's errors, and various types of scheduled walk-through inspections. As with conventional ETs, these recovery paths can easily be represented in HRA probability trees (Figure 8). In the case of annunciators or inspectors, the relevant failure limb is extended into two additional limbs: one failure limb and one success limb. The probability that the human responds successfully to the annunciator or that the inspector spots the operator's error is then fed back into the success path of the original tree. In the case of recovery by fellow team members, BHEPs are modified to CHEPs by considering the degree of dependency between the operator and one or more fellow workers who are in a position to notice the error. The effects of recovery factors can be determined by repeating the computations for total task failure.

In addition to considering error recovery factors, the analyst can choose to perform sensitivity analysis. One approach to sensitivity analysis is to identify the most probable errors on the tree, propose design modifications corresponding to those task elements, estimate the degree to which the corresponding HEPs would become reduced by virtue of these modifications, and evaluate the effect of these design interventions on the computation of the total task failure probability. The final step in THERP is to incorporate the results of the HRA into system risk assessments such as PRAs.

An obvious deficiency of THERP is its inability to handle human errors that have a more complex cognitive basis. Despite attempts to embellish THERP [e.g., through "sneak analysis" methods that may enable the analyst to identify decision making errors (Hahn and deVries, 1991)], THERP's underlying emphasis on decomposition and subsequent aggregation of individual actions has been questioned. For example, Hollnagel (1993) has argued that human reliability cannot be accounted for by considering "each action on its own" but rather, by considering "actions as a whole sequence," and has developed an alternative approach to HRA based on a modeling framework for predicting cognitive reliability that can also be used to support system risk assessments (Hollnagel, 1998). Although THERP's inability to adequately address more cognitively complex tasks and the underlying causality of human error tends to cast it as shallow, the insights concerning system operations acquired through THERP's attention to detail ultimately are likely to make it more useful than the numbers it makes available to quantitative risk assessments such as PRAs. In this respect, THERP shares many of the characteristics of PRAs: The quantitative products provided by PRAs are often considered to be less important than the ability for these risk assessments to identify deficiencies in design, provide a better understanding of interdependencies among systems and operations, and offer insights for improving procedures and operator training (Kumamoto and Henley, 1996).

5.3 SLIM-MAUD

The *success likelihood index methodology* (SLIM) represents another procedure for deriving HEPs (Embrey et al., 1984). In contrast to THERP, SLIM allows the analyst to focus on any human action or task. Consequently, this method can provide inputs into PRAs at various system levels; that is, the HEPs can reflect relatively low-level actions that cannot be further decomposed, as well as more broadly defined actions that encompass many of these lower-level actions. This increased flexibility, however, comes at the expense of a greatly reduced emphasis on task analysis and an increased reliance on subjective assessments.

SLIM assumes that the probability that a human will carry out a particular task or action successfully depends on the combined effects of a number of relevant PSFs. For each action under consideration, task domain experts are required to identify the relevant set of PSFs; assess the relative importance (or weights) of each of these PSFs with respect to the likelihood of some potential error mode associated with the action; and independent of this assessment, rate how good or bad each PSF actually is. Relative importance weights for the PSFs are derived by asking each analyst to assign a weight of 100 to the most important PSF, and then assign weights ranging from 0 to 100 to each of the remaining PSFs based on the importance of these PSFs relative to the one assigned the value of 100. Normalized weights are derived by dividing each weight by the sum of the weights for all the PSFs. The judges then rate each PSF on each action or task, with the lowest scale value indicating that the PSF is as poor as it is likely to be under real operating conditions, and the highest scale value indicating that the PSF is as good as it is likely to be in terms of promoting successful task performance. The likelihood of success for each human action is determined by summing the product of the normalized weights and ratings for each PSF, resulting in numbers (SLIs) that represent a scale of success likelihood.

The SLIs are useful in their own right. For example, if the actions under consideration represent alternative modes of response in an emergency scenario, the analyst may be interested in determining which types of responses are least or most likely to succeed. However, for the purpose of conducting PRAs, SLIM converts the SLIs to HEPs. An estimate of the HEP is derived using the following relationship:

$$\text{probability of success} = a \times \text{SLI} + b$$

where HEP is 1− the probability of success. To derive the two constants in this equation, the probabilities of success must be available for at least two tasks taken from the cluster of tasks for which the relevant set of PSFs was identified. However, methods exist for deriving HEPs even if information on such "reference" tasks is not available. Methods also exist for deriving upper and lower uncertainty bounds for these HEPs, which PRAs typically require.

Multiattribute utility decomposition (MAUD) provides a user-friendly computer-interactive environment for implementing SLIM. This feature ensures that many of the assumptions that are critical to the theoretical underpinnings of SLIM are met. For example, MAUD can determine if the ratings for the various PSFs by a given analyst are independent of one another and whether the relative importance weights elicited for the PSFs are consistent with the analyst's preferences. In addition, MAUD provides procedures for assisting the expert in identifying the relevant PSFs. Further details concerning SLIM–MAUD are provided in Embrey et al. (1984) and Kirwan (1994).

5.4 Human Error Data

As indicated in the discussion of THERP, fundamental data on HEPs can come from a variety of sources. Ideally, HEP data should derive from the relevant operating experience or at least from similar industrial experiences. However, as Kirwan (1994) notes, a number of problems are associated with collecting this type of quantitative HEP data. For example, many workers will be reluctant to report errors due to the threat of reprisals, and mechanisms for investigating errors are often nonexistent.

Even if these problems could be overcome, there are still other issues to contend with concerning the collection of useful HEP data. One problem is that errors that do not lead to a violation of a company's technical specifications or that are recovered almost immediately will probably not be reported. Also, data on errors associated with very low probability events, as in the execution of recovery procedures following an accident, may not be sufficiently available to produce reliable estimates and thus often require simulator studies for their generation. Finally, error reports are usually confined to the observable manifestations of an error (the external error modes). Without knowledge of the underlying cognitive processes or psychological mechanisms, errors that are in fact dissimilar (Table 1) may be aggregated. This would not only corrupt the HEP data but could also compromise error-reduction strategies.

In a study covering over 70 incidents in the British nuclear industry, it was possible to compile data on external error modes, PSFs, and psychological error mechanisms, and to derive 34 different HEPs (Kirwan et al., 1990), suggesting the possibility for collecting reasonably accurate operational-experience human error data. More typically, HEPs are derived from other sources, including expert judgments, laboratory experiments, and simulator studies. Table 9 presents examples of HEP data from several of these sources. Additional data on HEPs that include upper and lower uncertainty bounds and the effects of PSFs on nominal HEPs may be found in Swain and Guttmann (1983) and Gertman and Blackman (1994). More recently, Kirwan (1999) has reported on the construction of a HEP database in the UK referred to as CORE-DATA (computerized operator reliability and error database) for supporting HRA activities. CORE-DATA currently contains a large number of HEPs; its long-term objective is to apply its data to new industrial contexts through the development of extrapolation rules.

6 INCIDENT REPORTING SYSTEMS

6.1 Design, Data Collection, and Management Considerations

Information systems allow extensive data to be collected on incidents, accidents, and human errors, and thus afford excellent opportunities for organizations to learn. The distinction between accidents and incidents varies among authors and government regulatory agencies. Generally, accidents imply injury to persons or reasonable damage to property, whereas incidents usually involve the creation of hazardous conditions that if not recovered could lead to an accident. *Accidents* and *adverse events* are terms that are often used interchangeably, as are *incident*, *near miss*, and *close call.*

Capturing information on near misses is particularly advantageous. Depending on the work domain, near misses may occur hundreds of times more often than adverse events. If near misses are regarded as events that did not result in accidents by virtue of chance factors alone, the contexts surrounding near misses should be highly predictive of accidents. The reporting of near misses, especially in the form of short event descriptions or detailed anecdotal reports, would then provide a potentially rich set of data that could be used as a basis for proactive interventions. Moreover, fewer barriers exist in reporting them (Barach and Small, 2000). However, to anticipate hazardous scenarios and provide the proactive accident prevention function necessary for enabling organizations to improve continuously, *incident reporting systems* (IRSs) must be capable of identifying the underlying causes of the reported events.

The role of management is critical to the successful development and implementation of an IRS (CCPS, 1994). Management not only allocates the resources for developing and maintaining the system but can also influence the development of work cultures that may be resistive to the deployment of IRSs. In particular, organizations that have instituted "blame cultures" (Reason, 1997) are unlikely to advocate IRSs that emphasize underlying causes of errors, and workers in these organizations are unlikely to volunteer information to these systems. Ultimately, management's attitudes concerning human error causation will be reflected in the data that will be collected. The adoption of a system-induced perspective on human error that is consistent with Figure 1 would imply the need for an information system that emphasizes the collection of data on possible causal factors, including organizational and management policies responsible for creating the latent conditions for errors. Data on near misses would be viewed as indispensable for providing early warnings about how the interplay between human fallibility and situational contexts can penetrate barriers. System-based perspectives to human error are also conducive to a dynamic approach to data collection—if the methodology is proving inadequate in accounting for or anticipating human error, it will probably be modified (Figure 9).

Worker acceptance of an IRS that relies on voluntary reporting entails that the organization meet

Table 9 Examples of HEP Data Derived from Various Sources

Error	Probability
Data from Operational Plants	
1. Invalid address keyed into process-control computer This error occurred in a computer-controlled-batch chemical plant. When a valve sticks, or another malfunction occurs, the operator goes through a sequence on the computer which includes entering an address code for the component to be manipulated. The operator could, however, enter the wrong address (i.e., either an address for which there is no item, or the address for the wrong item); the HEP reflects the sum of these two alternative errors. There is a plant mimic available, prompt feedback is given of control actions, and the task occurs in normal operations.	0.007
2. Precision error: incorrect setting of chemical interface pressure In this event, an interface-pressure setting was set incorrectly, allowing an aqueous solution to pass into the stock tank, where it subsequently crystallized—which has a highly serious consequence. The error was caused largely by the failure on the part of the operator to be precise enough when setting the equipment.	0.03
3. Welders worked on wrong line Welders at a chemical plant worked on a vent line by mistake and holed a pipe.	0.04
4. Erroneous discharge of contaminants into the sea In this event, material was discharged into the sea erroneously, partly due to a communications failure across two shifts.	0.0007
5. Fuel-handling machine moved while still attached to a static fuel tank In this event, the fuel-handling machine in question, resembling a large overhead crane but with a very limited view, from the crane cab, of the flasks it carries, was moved by the operator while it was still in fact attached to a flask via flexible hoses, thus rupturing the hoses. This accident was in part caused by a communication failure across a shift break.	0.0005
6. Critical safety system not properly restored following maintenance In this event, a U.S. boiling-water-reactor (BWR) core-spray-pump system was left in an incorrect line-up configuration after testing. Testing is done by the operator in the CCR five times per year on five similar systems. This particular error occurred on the control switches on the CCR panels. The consequences are serious, since the effect is to disable a backup safety system.	0.0006
7. Operator works on wrong pump In this event, an operator on the plant was instructed to work on a pump in the west part of the plant but instead worked erroneously on the identical east plant.	0.03
8. Wrong fuel container moved In a supervised and heavily logged operation, the wrong fuel container was moved via the crane. The operator in the crane cab did not have a direct view of the containers but could only see them via a CCTV facility. The operator was, however, in communications with local operators who could see the containers directly.	0.0007
Data Derived from Ergonomics Experiments	
9. Human-recall performance with digital displays A six-digit sequence was presented for 2.6 seconds. The subject then had to write down the digit sequence in the intervening 10 seconds before the next sequence was presented. Seventy-two slides of six-digit sequences were shown to each subject. The error in question involved not writing down the correct sequence.	0.03
10. Inspectors' level of accuracy in spotting soldering defects in a complex system In a study of the capabilities of quality-control inspectors, the inspectors examined a complex unit with 1500 wires soldered to various terminals over a 3-hour period. Thirty defects had been placed in each unit, and the inspectors had to find all these defects, which were similar to the kinds of defects that they would find or look for every day.	0.2
11. Typing performance Each touch-typist in this experiment was instructed to type out a 1000-character piece of text as fast as possible without exceeding an error rate of 1%.	0.01
12. Network problem solving: a premature diagnosis Subjects were required to find the faulty component in a network of AND units. If two units feed into another unit, and one or both of the first two units are unhealthy, the third unit will read "unhealthy." However, the operator can only see which units are healthy or unhealthy at the end of the line of connected units, although he can also see how all the units are interconnected. Thus, the correct diagnosis involves determining which unit is unhealthy and is affecting the other units (only one unit is unhealthy in each network) and requires the operator to perform a type of fault diagnosis known as *backward chaining*. This task is very similar to a fault diagnosis for electrical maintenance panels. The number of units per network ranged from 16 to 24, always with four main "lines" leading to four final	0.07

Table 9 ***(continued)***

Error	Probability
output states (healthy or unhealthy). A *premature* diagnosis implied that the operator identified the faulty unit without first having carried out enough tests conclusively to determine which unit was faulty (irrespective of whether the premature diagnosis was correct or not: the task cannot afford the operator to make premature guesses).,	
13. Failure to carry out a one-step calculation correctly	0.01
14. Failure to carry out a seven- to 13-step calculation correctly	0.27
Simulator-Derived Data	
15. Emergency manual trip in a nuclear control room Prior to a fault appearing, the operator would be occupied with normal operations in a simulated control room. Initially, when a fault appeared, the operator was expected to try to control the fault, but it quickly became apparent that this was not possible, the operator was required instead to shut down (trip) the plant. The faults in question comprised a control-rod runout, a blower failure, a gas-temperature rise, and a coolant-flow fault. Tripping the plant required a single pushbutton activation. The fault rate in this scenario was 10 signals per hour (normally, it would have been on the order of 1 in 10,000 hours). The operator had only 30 to 90 seconds to respond by tripping the reactor, during which time the operator would have had to detect and diagnose the problem and then take action almost immediately.	0.2
16. Omission of a procedural step in a nuclear control room This HEP is based on a number of different scenarios, which were faced by shift teams in a full-scope nuclear power plant (NPP) simulation in the United States. The shift teams, all of whose members were being recertified as NPP operators, were required to deal with a number of emergency scenarios.	0.03
17. Selection of wrong control (discrimination by label only) This HEP, which was derived from a number of NPP simulator scenarios, was based on 20 incorrect (unrecovered) selections from out of a total of 11,490 opportunities for control selection.	0.002
18. Selection of wrong control (functionally grouped) As above, but this time the HEP is based on only four unrecovered errors out of 27,055 opportunities for error.	0.0002
19. Equipment turned in wrong direction As above, based on the unrecovered errors again, and with equipment that does not violate a population stereotype (i.e., with normal, expected turning conventions).	0.00002

Source: Kirwan (1994).

three requirements: exact a minimal use of blame; ensure freedom from the threat of reprisals, and provide feedback indicating that the system is being used to affect positive changes that can benefit all stakeholders. Accordingly, workers would probably not report the occurrence of accidental damage to an unforgiving management and would discontinue voluntarily offering information on near misses if insights gained from intervention strategies are not shared (CCPS, 1994). It is therefore essential that reporters of information perceive IRSs as error management or learning tools and not as disciplinary instruments.

In addition to these fundamental requirements, two other issues need to be considered. First, consistent with user-centered design principles (Nielsen, 1995), potential users of the system should be involved in its design and implementation as they would with any newly designed (or redesigned) product, although for very large populations of potential users this may not be practical. Second, effective training is critical to the system's usefulness and usability. When human errors, near misses, or incidents occur, the people who are responsible for their reporting and investigation need to be capable of addressing in detail all considerations related to human fallibility, context, and barriers that affect the incident. Thus, training may be required for recognizing that an incident has in fact occurred and for providing full descriptions of the event. Training would also be necessary for ensuring that these data are input correctly into the information system and for verifying that the system's knowledge base adequately supports the representation of this information. Analysts would need training on applying the system's tools, including the use of any modeling frameworks for analyzing causality of human error and on interpreting the results of these application tools. They would also need training on generating summary reports and recommendations and on making modifications to the system's database and inferential tools if the input data imply the need for such adjustments. Access control would also need to be addressed. For each category of system user (e.g., manager, human factors analyst, employee) a *reading authority* (who is allowed to retrieve information from the system) and a *writing authority* (who is allowed to update the database) need to be specified.

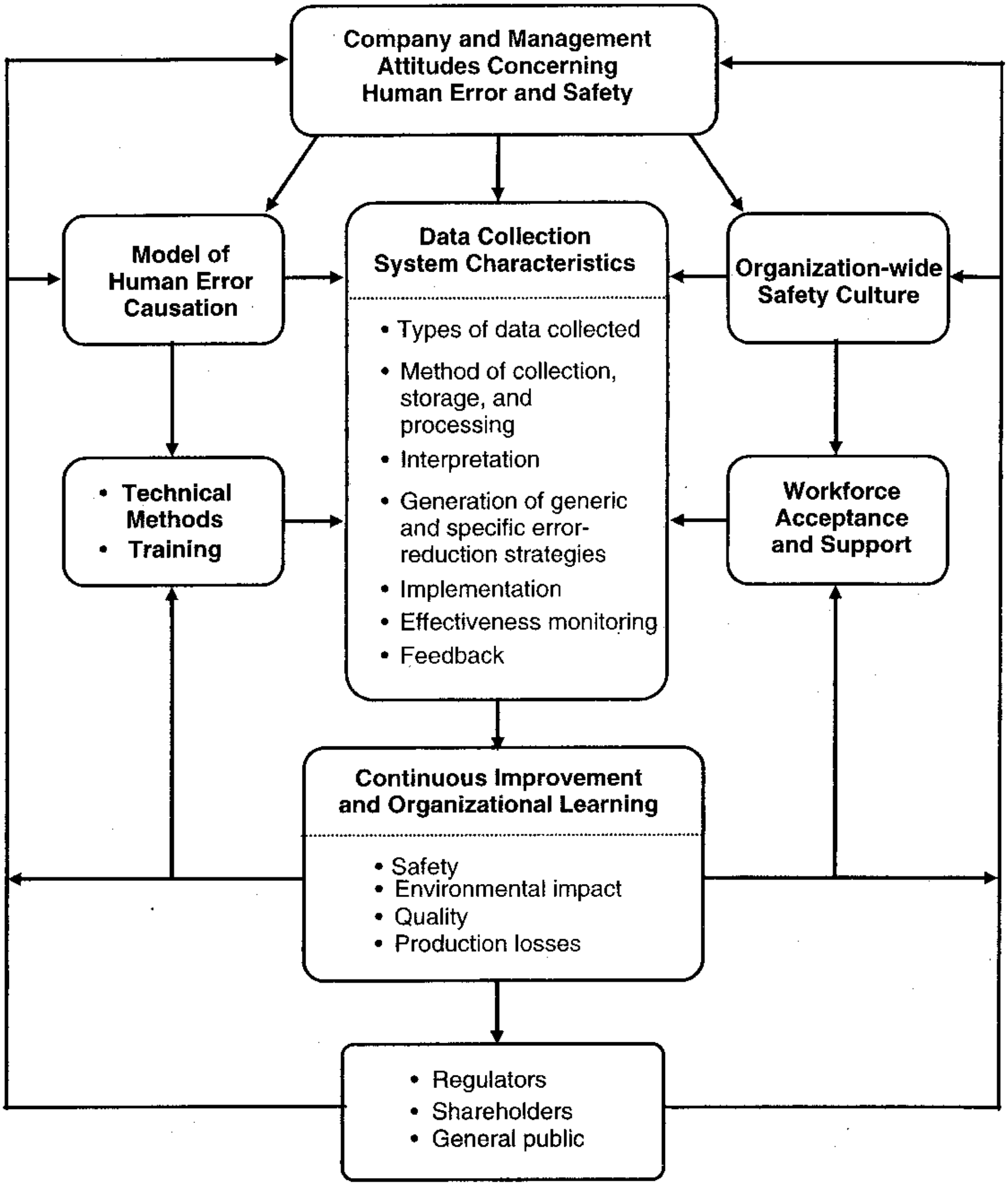

Company and Management Attitudes Concerning Human Error and Safety
Model of Human Error Causation
Data Collection System Characteristics
• Types of data collected
• Method of collection, storage, and processing
• Interpretation
• Generation of generic and specific error-reduction strategies
• Implementation
• Effectiveness monitoring
• Feedback
Organization-wide Safety Culture
• Technical Methods
• Training
Workforce Acceptance and Support
Continuous Improvement and Organizational Learning
• Safety
• Environmental impact
• Quality
• Production losses
• Regulators
• Shareholders
• General public

Figure 9 Data collection system for error management. (Adapted from CCPS, 1994. Copyright 1994 by the American Institute of Chemical Engineers, and reproduced by permission of AIChE.)

Data for input into IRSs can be of two types: quantitative data, which lend themselves more easily to coding and classification, and qualitative data in the form of free-text descriptions. Kjellén (2000) has specified the basic requirements for a safety information system in terms of data collection, distribution and presentation of information, and overall information system attributes. To meet data collection requirements, the input data need to be reliable (if the analysis were to be repeated, it should produce similar results), accurate, and provide adequate coverage (e.g., on organizational and human factors/ergonomics issues) needed for exercising efficient control. Foremost in the distribution and presentation of information is the need for relevant information. *Relevance* will depend on how the system will be used. If the objective is to analyze statistics on accidents in order to assess trends, a limited set of data on each accident or near miss would be sufficient and the nature of these data can often be specified in advance. However, suppose that the user is interested in querying the system regarding the degree to which new technology and communication issues have been joint factors in incidents involving errors of omission. In this case, the relevance will be decided by the coverage. Generally, the inability to derive satisfactory answers to specific questions will signal the need for modifications of the system.

In addition to relevance, the information should be comprehensible and easy to survey; otherwise, its use will be restricted to highly trained analysts, prompting high-level management to view the system with suspicion. Overall, the information system should promote involvement between management and employees, thus fostering organizational learning. Finally, the system should be cost-efficient. As in most cost–benefit analyses, costs will be much easier to assess than benefits. Investment, operations, and maintenance costs are relatively straightforward to determine, as are potential benefits resulting from cost reductions associated with the handling, storing, and distribution of various safety-related documents. Benefits associated with reductions in adverse outcomes such as accidents, production delays, and reduced qualities are generally much more difficult to assess.

In searching the database, the user may restrict the search to events that meet criteria defined on one of the standard four (nominal, ordinal, interval, or ratio) scales of measurement (e.g., find all near misses involving workers with less than six months of experience) or to events that include keywords in free-text descriptions (e.g., find all near misses of radiation overexposure that resulted in disruptions to production schedules). Data entered and coded based on standard forms of measurement are relatively easy to manage, whereas data that have been documented and stored in unstructured free-text descriptions may require intelligent software agents for analysis and interpretation. All information searches, however, afford the possibility for type I errors (wanted data that are not found) and type II errors (unwanted data identified as hits).

6.2 Historical Antecedent

An idea related to IRSs, that of the modest suggestion box, has been around for hundreds of years. Both IRSs and suggestion programs are tools designed to capture problem-related data from interested parties regarding the operations of an organization, and are deployed by organizations in order to learn about and improve themselves. One of the earliest suggestion programs, implemented by the British Navy in 1770 (Robinson and Stern, 1998), was motivated by the recognition that persons within the organization should have a way of speaking out without fear of reprisals. The first suggestion box was implemented in the Scottish firm William Denny & Brothers in 1880, and the first U.S. company to implement a companywide suggestion program was National Cash Register in 1892. The suggestion program gained rapid acceptance following World War II, when it was adapted by quality initiatives to meet various objectives, such as safety (Turrell, 2002). At the Toyota Motor Corporation, the suggestion program is part of the Kaizen or "continuous improvement" approach to manufacturing and represents an extremely important feature of the Toyota production system. Implemented in 1951, it took nine years to achieve a 20% participation rate. In 1999, data from the Toyota Motor manufacturing plant in Kentucky indicated that 5048 of 7800 employees contributed 151,327 ideas into the system and that nearly all were implemented (Leech, 2004), resulting in $41.5 million in savings. Unquestionably, the benefits that can potentially be accrued from constructive use of worker feedback can have a powerful impact on an organization's effectiveness and are the basis for the appeal of IRSs in industry.

6.3 The Aviation Safety Reporting System

The Aviation Safety Reporting System (ASRS) was developed in 1976 by the Federal Aviation Administration (FAA) in conjunction with the National Aeronautics and Space Administration (NASA). Many significant improvements in aviation practices have since been attributed to the ASRS, and these improvements have largely accounted for the promotion and development of IRSs in other work domains (Table 10).

The ASRS's mission is threefold: to identify deficiencies and discrepancies in the National Aviation System (NAS), to support policy formulation and planning for the NAS, and to collect human performance data and strengthen research in the aviation domain. All pilots, air traffic controllers, flight attendants, mechanics, ground personnel, and other personnel associated with aviation operations can submit confidential reports if they have been involved in or observed any incident or situation that could have a potential effect on aviation safety. Preaddressed postage-free report forms are available online and are submitted to the ASRS via the U.S. Postal Service. However, unlike other systems, the ASRS presently is not equipped to handle online submissions of information. The ASRS database can be queried by accessing its Internet site (http://asrs.arc.nasa.gov), and is also available on CD-ROM.

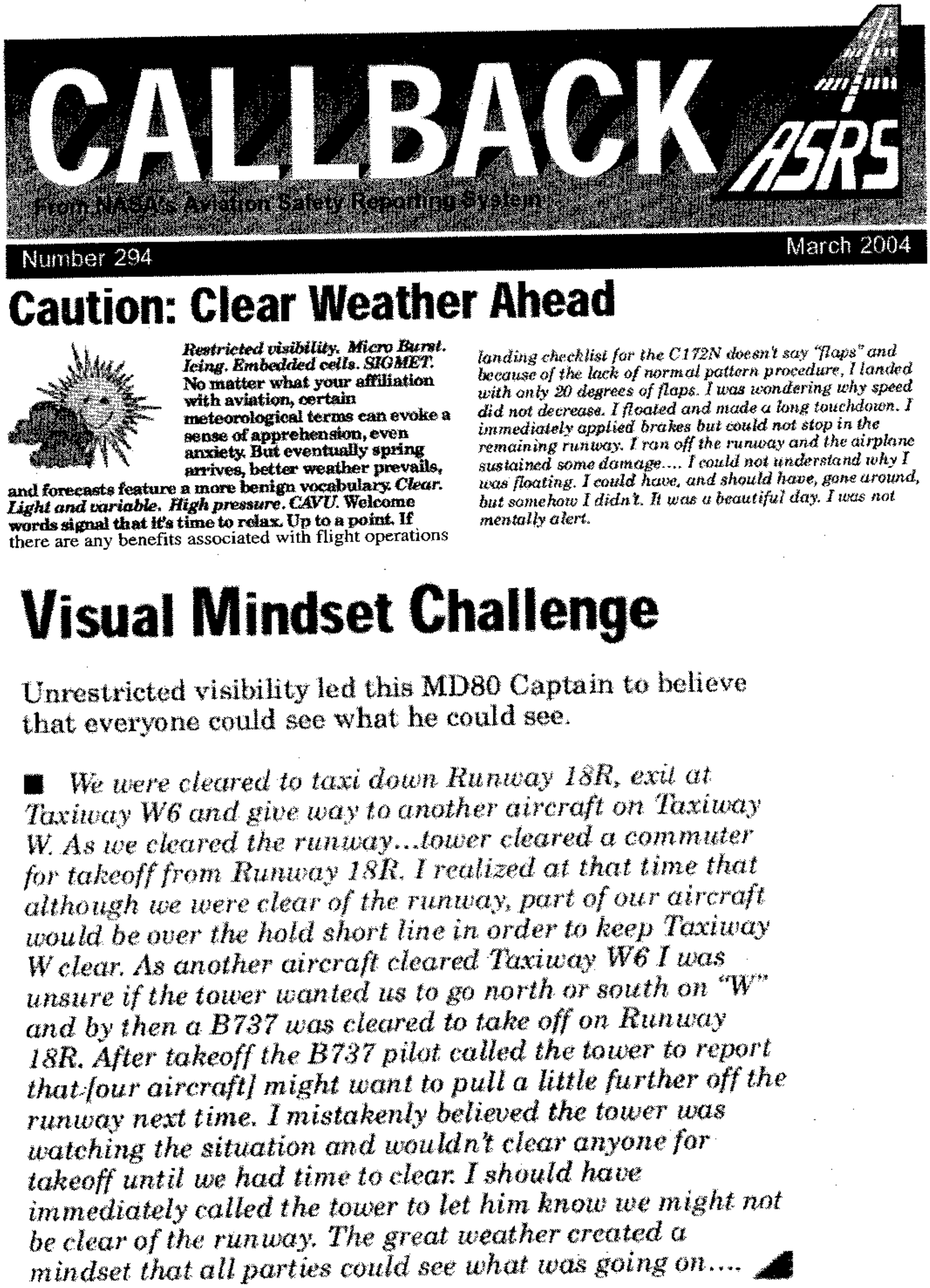
ASRS reports are processed in two stages by a group of analysts composed of experienced pilots and air traffic controllers. In the first stage, each report is read by at least two analysts who identify incidents and situations requiring immediate attention. Alerting messages are then drafted and sent to the appropriate group. In the second stage, analysts classify the reports and assess causes of the incident. Their analyses and the information contained in the reports are then incorporated into the ASRS database. The database consists of the narratives submitted by each reporter and coded information that is used for information retrieval and statistical analysis procedures.

Several provisions exist for disseminating ASRS outputs. These include alerting messages that are sent out in response to immediate and hazardous situations, the *CALLBACK* safety bulletin, which is a monthly publication containing excerpts of incident report narratives and added comments (Figure 10), and the ASRS *Directline*, which is published to meet the needs of airline operators and flight crews. In addition, in response to database search requests, ASRS staff communicates with the FAA and the National Transportation Safety Board (NTSB) on an institutional level in support of various tasks, such as accident investigations, and conducts and publishes research related primarily to human performance issues.

Table 10 Attributes of Incident Reporting Systems

Reporting System	Ownership	Regulatory	Mandatory	Voluntary	Anonymous	Confidential	Narrative	Immunity	Threshold	Feedback
Aviation safety reporting system	Federally funded, administered by NASA	Yes	No	Yes	After filed	Yes	Yes	Yes	All nonaccidents	Yes (*Callback*)
Aviation safety airways program	American Airlines	No	No	Yes	No	Yes	Yes	No	All noncrashes	Yes
Airline Pilots Association	FAA in with private pilot association	No	No	Yes	No	Yes	Yes	No	All incidents	Yes
British Airways safety information system										
Air safety report	British Airways	No	Yes	No	No	Yes	Yes	No	Safety-related events	Yes (*Flywise)*
Confidential human factors reporting program	British Airways	No	No	Yes	No	Yes	No, but can expand	No	Human factor data	Yes
Special event search and master analysis	British Airways	Yes	Yes	No	Yes	Yes	N/A	Yes	Monitors fight data records	Yes
Human factors failure analysis classification system	U.S. navy and marines	Yes	Yes	No	No	No	Yes	No	All crashes	Yes
NASA	Federal	Yes	Yes	No	No	Yes	Yes	No	All safety events	Yes
Prevention and recovery information system for monitoring and analysis	Institutional	No	No	Yes	Yes	Yes	Yes	No	Accidents and near misses	Yes
Human factors information systems	Federal with private input (INPO)	Yes	No	Yes	No	Yes	Yes	Yes	Human factor issues related to nuclear safety	Yes
NRC allegations systems process	Federal	Yes	No	Yes	No	Yes	Yes	Yes	All safety concerns	Yes
Diagnostic misadministration reports-regulatory information distribution system	Federal, nuclear regulatory control	Yes	Yes	No	No, patient ID is needed	No	Yes	?	All misadministration	Yes

Source: Barach and small (2000).



CALLBACK ASRS

From NASA's Aviation Safety Reporting System

Number 294 March 2004

Caution: Clear Weather Ahead

Restricted visibility. Micro Burst. Icing. Embedded cells. SIGMET. **No matter what your affiliation with aviation, certain meteorological terms can evoke a sense of apprehension, even anxiety. But eventually spring arrives, better weather prevails, and forecasts feature a more benign vocabulary.** ***Clear. Light and variable. High pressure. CAVU.*** **Welcome words signal that it's time to relax. Up to a point. If** there are any benefits associated with flight operations

landing checklist for the C172N doesn't say "flaps" and because of the lack of normal pattern procedure, I landed with only 20 degrees of flaps. I was wondering why speed did not decrease. I floated and made a long touchdown. I immediately applied brakes but could not stop in the remaining runway. I ran off the runway and the airplane sustained some damage.... I could not understand why I was floating. I could have, and should have, gone around, but somehow I didn't. It was a beautiful day. I was not mentally alert.

Visual Mindset Challenge

Unrestricted visibility led this MD80 Captain to believe that everyone could see what he could see.

■ *We were cleared to taxi down Runway 18R, exit at Taxiway W6 and give way to another aircraft on Taxiway W. As we cleared the runway...tower cleared a commuter for takeoff from Runway 18R. I realized at that time that although we were clear of the runway, part of our aircraft would be over the hold short line in order to keep Taxiway W clear. As another aircraft cleared Taxiway W6 I was unsure if the tower wanted us to go north or south on "W" and by then a B737 was cleared to take off on Runway 18R. After takeoff the B737 pilot called the tower to report that [our aircraft] might want to pull a little further off the runway next time. I mistakenly believed the tower was watching the situation and wouldn't clear anyone for takeoff until we had time to clear. I should have immediately called the tower to let him know we might not be clear of the runway. The great weather created a mindset that all parties could see what was going on....*

Figure 10 Excerpt of an incident report narrative from the *CALLBACK* bulletin.

Many major U.S. airlines have their own internal programs for tracking human errors, especially among pilots, and these programs usually rely on some form of IRS. The Aviation Safety Action Partnership (ASAP), an American Airlines program, is somewhat unique in that it collaborates with the FAA to determine how pilot errors should be handled. The event-review team includes an FAA official, company managers, and representatives from the Allied Pilots Association (the pilots' union). Typically, if the FAA determines that a pilot error has occurred, it issues a citation with penalties ranging from warning letters to license revocation. However, American Airlines pilots who file ASAP reports are assured that the FAA will exact no punishment, or less severe punishment, as long as the error was unintentional. It has been estimated that without ASAP the FAA would be aware of fewer than 1% of the errors of American Airline pilots (Kaye, 1999b).

In contrast to these fairly conventional industry-specific IRSs, United Airlines has adopted a more sophisticated approach to dealing with pilot error.

Its flight operations quality assurance program uses optical recorders in most of its daily flights to reduce pilot error by capturing a pilot's every move electronically (Kaye, 1999c). These disks are later analyzed by computer, and if something wrong, dangerous, or outside normal operating procedure is identified, a team of 10 United Airlines pilots examines the problem and determines a course of action. For example, if a proficiency issue is identified, the team can authorize training for that pilot. The optical recorders could also be connected to operating systems other than the cockpit. For instance, when linked to its maintenance systems, it enabled United Airlines to discover that some internal engine parts were cracking from too much heat. Electronic monitoring presumes relinquishing privacy; thus, acceptance of the program will require that workers acknowledge the possibility that more of their mistakes can become corrected.

6.4 Medical Incident Reporting Systems

The medical industry is currently struggling with what has been termed an epidemic of adverse events stemming from medical error. (This industry defines an adverse event as an injury or death resulting from medical management, and medical error as the failure of a planned action to be completed as intended.) Taking a cue from other complex high-risk industries such as nuclear power and chemical processing, the health care industry is increasingly considering, developing, and deploying IRSs to deal with patient safety concerns and related issues. Despite acknowledgment by the Institute of Medicine that there are an enormous number of preventable injuries to patients (Kohn et al., 1999), implementing IRSs in the health care industry has lagged behind other industries and for good reason. Compared to other industries, the health care industry interacts with the public on a highly personal basis, and protecting reports on near misses, incidents, and accidents is likely to be met with resistance from a public that especially in the United States, is entrenched in a culture of litigation and that is seeing an increasing part of their income being allocated to health care costs (Section 1.1). Collecting reports on medical error that are anonymous may not appease the public—amnesty of unsafe acts that lead to near misses and adverse events would probably not go over very well. Interestingly, medical IRSs have been successful in gaining acceptance in Australia and New Zealand, where legal protection for those who report events has been enforced (Rosenthal et al., 2001).

Not surprisingly, standardization of definitions of errors, near misses, and adverse events in the medical industry, which is fundamental to the industry's ability to gather information, learn about patient safety, and institute intervention strategies, has been difficult to establish. It is also questionable whether a true safety culture that supports IRSs exists in the medical industry. Despite these issues, there have been a number of successful implementations of IRSs in the health care industry, in particular in transfusion medicine, intensive care, anesthesia, occupational medicine, and pharmacy. One example is the Veteran's Administration Patient Safety Reporting System (PSRS), which developed out of an agreement in 2000 between NASA and the Department of Veteran's Affairs (VA). The PSRS allows all VA medical facility staff to report voluntarily any events and concerns related to patient safety confidentially without being subject to reprisals. The types of events that can be reported include close calls (i.e., near misses), unexpected situations involving death, physical or psychological injury of a patient or employee, and lessons learned related to patient safety. Ultimately, the information is made available through alerts, publications such as the Patient Safety Bulletin, and research studies. Although still in use, the PSRS now serves as a complement to a more recent reporting system being operated by the VA that utilizes a root-cause analysis methodology (Section 10.2) for analyzing adverse events and near misses, and provides strategies for decreasing the likelihood of the event's reoccurrence.

Another example of a medical IRS is the Anesthesia Critical Incident Reporting System (CIRS) operated by the Department of Anesthesia at the University of Basel in Switzerland. Using a Web-based interface, contributors worldwide can anonymously report information on incidents in anesthesia practice and review information collected on those incidents. The CIRS IRS defines a *reportable event* as "an event under anesthetic care which has the potential to lead to an undesired outcome if left to progress." Contributors can also report events resulting from team interactions. The design of this system was based on the experiences of the Australian AIMS study, another influential IRS for reporting anesthesia incidents.

As of this writing, the U.S. Senate has proposed a bill that would set up a confidential, voluntary system for reporting medical errors in hospitals without fear of litigation. The goal of the bill, which is pending committee review and action, is to encourage health care providers to report errors so they can be analyzed by patient safety organizations for the purpose of producing better procedures and safety protocols that could improve the quality of care. Notably, in his statement supporting the passage of this bill, Donald Palmisano, the immediate past president of the American Medical Association, stated that "the Aviation Safety Reporting System serves as a successful model for this system."

6.5 Limitations of Incident Reporting Systems

Some IRSs, by virtue of their inability to cope with the vast number of incidents in their databases, have apparently become "victims of their own success" (Johnson, 2002). The Federal Aviation Administration's (FAA's) ASRS and the Food and Drug Administration's MedWatch Reporting System (designed to gather data on regulated, marketed medical products, including prescription drugs, specialized nutritional products, and medical devices) both contain over a half a million incidents. Because their database technologies were not designed to manage this magnitude of data, users who query these systems are having trouble extracting useful information and often fail to identify important cases. This is particularly true of the many IRSs that rely on *relational database* technology. In these systems, each incident is stored as a record and incident identifiers are used to link similar records in response to user queries. Relational database techniques, however, do not adapt well to changes in the nature of incident reporting or in the models of incident causation. Also, different organizations in the same industry tend to classify events differently, which reduces the benefits of drawing on the experiences of IRSs across different organizations. It can also be extremely difficult for people who were not involved in the coding and classification process to develop appropriate queries (Johnson, 2002).

Problems with IRSs can also arise when large numbers of reports on minor incidents are stored. These database systems may then begin to drift toward reporting information on quasi-incidents and precursors of quasi-incidents, which may not necessarily provide the IRS with increased predictive capability (Amalberti, 2001). As stated by Amalberti: "The result is a bloated and costly reporting system with not necessarily better predictability, but where everything can be found; this system is chronically diverted from its true calling (safety) to serve literary or technical causes. When a specific point needs to be proved, it is (always) possible to find confirming elements in these extra-large databases" (p. 113). There is, however, a counterargument to this view: that in the absence of a sufficient number of true incidents, the judicious examination of quasi-incidents may reveal vulnerabilities within the system that would normally be concealed. In this regard, exploiting the potential of quasi-incidents in IRSs suggests the possibility for a proactive capability that may indeed reflect the existence of a highly evolved safety culture.

There is a drift of a different sort that would be advantageous to catalog, but unfortunately is not amenable to capture by the current state-of-the-art in incident reporting. These drifts reflect the various adaptations by an organization's constituents to the external pressures and conflicting goals to which they are continuously subjected (Dekker, 2005). *Drifting into failure* may occur, for instance, when a worker confronts increasingly scarce resources while under pressure to meet higher production standards. If the adaptive responses by the worker to these demands gradually become absorbed into the organization's definition of normal work operations (Section 8), work contexts that may be linked to system failures are unlikely to be reported. The intricate, incremental, and transparent nature of the adaptive processes underlying these drifts is manifest at both the horizontal and vertical levels of an organization. Left unchecked, aggregation of these drifts seals an organization's fate by effectively excluding the possibility for proactive risk management solutions. In the case of the accident in Bhopal (Casey, 1993), these drifts were personified at all levels of the responsible organization. Although IRSs can, in theory, monitor these types of drifts, to do so these systems may need to be driven by new models of organizational dynamics and armed with new levels of intelligence (Dekker, 2005).

A much more fundamental problem with IRSs is the difficulty in assuring anonymity to reporters, especially in smaller organizations. Although most IRSs are confidential, anonymity is more conducive to obtaining disclosures of incidents. Unfortunately, anonymity precludes the possibility for follow-up interviews, which are often necessary for clarifying reported information (Reason, 1997).

Being able to follow up interviews, however, does not always resolve problems contained in reports. Gaps in time between the submission of a report and the elicitation of additional contextual information can result in important details being forgotten or confused, especially if one considers the many forms of bias that can affect eyewitness testimony (Table 11). Biases that can affect reporters of incidents can also affect the teams of people (i.e., analysts) that large-scale IRSs often employ to analyze and classify the reports. For example, there is evidence that persons who have received previous training in human factors are more likely to diagnose human factors issues in incident reports than persons who have not received this type of training (Lekberg, 1997).

Variability among analysts can also derive from the confusion that arises when IRSs employ classification schemes for incidents that are based on detailed taxonomies. Difficulty in discriminating between the various terms in the taxonomy may result in low recall systems, whereby some analysts fail to identify potentially similar incidents. In general, concerns associated with interanalyst reliability stemming from bias and differences in analysts' abilities can impede an organization's ability to learn. More specifically, limitations in analysts' abilities to interpret causal events reduces the capability for organizations to draw important conclusions from incidents, and analyst bias can lead to organizations using IRSs for supporting existing preconceptions concerning human error and safety. As alluded to earlier, training all analysts to the same standard, although a resource-intensive proposition for large organizations, is necessary for minimizing variability associated with inferring causality.

Although there are no software solutions to all these problems, a number of recommendations discussed by Johnson (2002) deserve consideration. For example, for IRSs that are confidential but not anonymous, computer-assisted interviewing techniques can mitigate some of the problems associated with follow-up elicitations of contextual details from reporters. By relying on frames and scripts that are selected in response to information from the user, these techniques can ensure that particular questions are asked in particular situations, thus reducing interanalyst biases stemming from the use of different interview approaches. The success of these approaches, however, depends on ensuring that the dialogue is appropriate for the situation that the reporter is being asked to address. Information-retrieval engines that are the basis for Web search also offer promise, due

Table 11 Forms of Eyewitness Testimony Biases in Reporting

- *Confidence bias:* arises when witnesses unwittingly place the greatest store in their colleagues who express the greatest confidence in their view of an incident. Previous work into eyewitness testimonies and expert judgments has shown that it may be better to place greatest trust in those who do not exhibit this form of overconfidence (Johnson, 2003).
- *Hindsight bias:* arises when witnesses criticize individuals and groups on the basis of information that may not have been available at the time of an incident.
- *Judgment bias:* arises when witnesses perceive the need to reach conclusions about the cause of an incident. The quality of the analysis is less important than the need to make a decision.
- *Political bias:* arises when a judgment or hypothesis from a high-status member commands influence because others respect that status rather than the value of the judgment itself. This can be paraphrased as "pressure from above."
- *Sponsor bias:* arises when a witness testimony can affect indirectly the prosperity or reputation of the organization they manage or for which they are responsible. This can be paraphrased as "pressure from below."
- *Professional bias:* arises when witnesses may be excluded from the society of their colleagues if they submit a report. This can be paraphrased as "pressure from beside."
- *Recognition bias:* arises when witnesses have a limited vocabulary of causal factors. They actively attempt to make any incident "fit" with one of those factors, irrespective of the complexity of the circumstances that characterize the incident.
- *Confirmation bias:* arises when witnesses attempt to make their evidence confirm an initial hypothesis.
- *Frequency bias:* occurs when witnesses become familiar with particular causal factors because they are observed most often. Any subsequent incident is therefore likely to be classified according to one of these common categories irrespective of whether an incident is actually caused by those factors.
- *Recency bias:* occurs when a witness is heavily influenced by previous incidents.
- *Weapon bias:* occurs when witnesses become fixated on the more "sensational" causes of an incident. For example, they may focus on the driver behavior that led to a collision rather than the failure of a safety belt to prevent injury to the driver.

Source: Adapted from Johnson (2002).

to their flexibility in exploiting semantic information about the relationships between terms or phrases that are contained in a user's query and in the reports. In some instances, these search techniques have been integrated with relational databases in order to capitalize on fields previously encoded into the database. However, the integration of these techniques cannot assure users that their queries will find similar incidents (i.e., the precision may be low), or as the large results lists that are typically generated from Web-based searches imply, return almost every report in the system (i.e., recall may be too high). Alternatives to relational databases and information retrieval techniques that have been suggested include conversational case-based reasoning, where the user must answer a number of questions in order to obtain information concerning incidents of interest. The possibility also exists for determining differences among analysts in the patterns of their searches, and thus insights into their potential biases, by tracing their interactions with these systems (Johnson, 2003).

Finally, a very different type of concern with IRSs arises when these systems are used as a basis for quantitative human error applications. In these situations, the voluntary nature of the reporting may invalidate the data that are used for deriving error likelihoods (Thomas and Helmreich, 2002). From a probabilistic risk assessment (Section 5.1) and risk management perspective, this issue can undermine decisions regarding allocating resources for resolving human errors: Which errors do you attempt to remediate if it is unclear how often the errors are occurring?

7 AUTOMATION AND HUMAN ERROR

7.1 Human Factors Considerations in Automation

Innovations in technology will always occur and will bring with them new ways of performing tasks and doing work. Whether the technology completely eliminates the need for the human to perform a task or results in new ways of performing tasks through automation of selective task functions, the human's tasks will probably become reconfigured (Chapter 60). The human is especially vulnerable when adapting to new technology. During this period, knowledge concerning the technology and the impact it may have when integrated into task activities is relatively unsophisticated and biases deriving from previous work routines are still influential.

Automating tasks or system functions by replacing the human's sensing, planning, decision making, or manual activities with computer-based technology often requires making allocation of function decisions—that is, deciding which functions to assign to the human and which to delegate to automatic control (Sharit, 1997). Because these decisions ultimately can have an impact on the propensity for human error, consideration may also need to be given to the level of automation to be incorporated into the system (Parasuraman et al., 2000; Kaber and Endsley, 2004). Higher levels imply that automation will assume greater autonomy in decision making and control. The primary concern with technology-centered systems is that they deprive themselves of the benefits deriving from the human's ability to anticipate, search for, and discern relevant data based on the current context; make generalizations and inferences based on past experience; and modify activities based on changing constraints. Determining the optimal level of automation, however, is a daunting task for the designer.

While levels of automation somewhere between the lowest and highest levels may be the most effective way to exploit the combined capabilities of both the automation and the human, identifying an ideal level of automation is complicated by the need also to account for the consequences of human error and system failures (Moray et al., 2000).

Many of the direct benefits of automation are accompanied by indirect benefits in the form of error reduction. For example, the traffic alert and collision avoidance system in aviation that assesses airspace for nearby traffic and warns the pilot if there is a potential for collision can overcome human sensory limitations, and robotic assembly cells in manufacturing can minimize fatigue-induced human errors. Generally, reducing human physical and cognitive workload enables the human to attend to other higher-level cognitive activities, such as the adoption of strategies for improving system performance. Reckless design strategies, however, that automate functions based solely on technical feasibility can often lead to a number of problems (Bainbridge, 1987). For instance, manual and cognitive skills that are no longer used due to the presence of automation will deteriorate, jeopardizing the system during times when human intervention is required. Situations requiring rapid diagnosis that rely on the human having available or being able quickly to construct an appropriate mental model will thus impose higher WM demands on humans who are no longer actively involved in system operations. The human may also need to allocate significant attention to monitoring the automation, which is a task humans do not perform well. These problems are due largely to the capability for automation to insulate the human from the process, and are best handled through training that emphasizes ample hands-on simulation exercises encompassing varied scenarios. The important lesson learned is that "disinvolvement can create more work rather than less, and produce a greater error potential" (Dekker, 2005, p. 165).

Automation can also be clumsy for the human to interact with, making it difficult to program, monitor, or verify, especially during periods of high workload. A possible consequence of clumsy automation is that it "tunes out small errors and creates opportunities for larger ones" (Weiner, 1985) by virtue of its complex connections to, and control of important systems. Automation has also been associated with *mode errors*, a type of mistake in which the human acts based on the assumption that the system is in a particular mode of operation (either because the available data support this premise or because the human instructed the system to adopt that mode) when in fact it is in a different mode. In these situations, unanticipated consequences may result if the system remains capable of accommodating the human's actions. The tendency for a system to mask its operational mode represents just one of the many ways that automation can disrupt situation awareness.

More generally, when the logic governing the automation is complex and not fully understood by the human, the actions taken by automatic systems may appear confusing. In these situations, the human's tendency for partial matching and biased assessments (Section 3.2) could lead to the use of an inappropriate rule for explaining the behavior of the system—a mistake that in the face of properly functioning automation could have adverse consequences. These forms of human–automation interaction have been examined in detail in flight deck operations in the cockpit and have been termed *automation surprises* (Woods et al., 1997). Training that allows the human to explore the various functions of the automation under a wide range of system or device states can help reduce some of these problems. However, it is also essential that designers work with users of automation to ensure that the user is informed about what the automation is doing and the basis for why it is doing it. In the past, slips and mistakes by flight crews tended to be errors of commission. With automation, errors of omission have become more common, whereby problems are not perceived and corrective interventions are not made in a timely fashion.

Another important consideration is *mistrust* of automation, which can develop when the performance of automatic systems or subsystems is perceived to be unreliable or uncertain (Lee and Moray, 1994). Lee and See (2004) have defined *trust* as an attitude or expectancy regarding the likelihood that someone or something will help the person achieve his or her goal in situations characterized by uncertainty and vulnerability. As these authors have pointed out, many parallels exist between the trust that we gain in other people and the trust we acquire in complex technology, and as in our interactions with other people, we tend to rely on automation we trust and reject automation we do not trust. Mistrust of automation can provide new opportunities for errors, as when the human decides to assume manual control of a system or decision-making responsibilities that may be ill-advised under the current conditions.

Like many decisions people make, the decision to rely on automation can be strongly influenced by emotions. Consequently, even if the automation is performing well, the person's trust in it may become undermined if its responses are not consistent with expectations (Rasmussen et al., 1994). Mistrust of automation can also lead to its disuse, which impedes the development of knowledge concerning the system's capabilities and thus further increases the tendency for mistrust and human error. Overreliance on automation can also lead to errors in those unlikely but still possible circumstances in which the automation is malfunctioning, or when it encounters inputs or situations unanticipated in its design that the human believes it was programmed to handle.

Lee and See (2004) have developed a conceptual model of the processes governing trust and its effect on reliance that is based on a dynamic interaction among the following factors: the human, organizational, cultural, and work contexts; the automation; and the human–automation interface. As a framework for guiding the creation of appropriate trust in automation, their model suggests that the algorithms governing the automation need to be made more transparent to the user, that the interface should provide information regarding the capabilities of the automation in a format that is easily understandable, and that training should address the varieties of situations that can affect the capabilities of the automation.

Organizational and work culture influences also need to be considered. If automation is imposed on workers, especially in the absence of a good rationale regarding its purpose or how human–automation interaction may enhance the work experience or improve the potential for job enrichment, the integrity and meaningfulness of work may become threatened, resulting in work cultures that promote unproductive and possibly dangerous behavioral strategies. Finally, as the work of Cao and Taylor (2004) described below suggests, the adverse effects that interacting with complex technology can have on team communication may require the need to address the concept of *meta-trust*, the trust people have that other people's trust in automation is appropriate (Lee and See, 2004).

7.2 Examples of Human Error in Commercial Aviation

The cockpits of commercial airliners contain numerous automated systems. Central among these systems is the flight management system (FMS). The FMS can be programmed to follow an assigned flight plan route, allowing a plane to navigate itself to a series of checkpoints and providing the estimated time and distance to these checkpoints. It can also determine speed and power settings that optimize fuel consumption, prevent the plane from descending below an altitude restriction, and display navigational information. Working in conjunction with the FMS is the autopilot, which allows the plane to assume and maintain a specific heading, level off at an assigned altitude, or climb or descend at a specific rate, and an auto-throttle system, which sets the throttles for specific airspeeds. In addition, the traffic alert and collision avoidance system notifies pilots about potential collisions with other aircraft and provides instructions on how to avoid that aircraft, the stormscope warns pilots when severe weather lies ahead, and the wind shear system allows pilots to detect wind shear during takeoff and approach to landing. The pilot can also employ an automatic landing system. These automatic systems have the potential to reduce pilot workload significantly and thus enhance safety. However, they can perform so many functions that pilots can lose sight of where they are or what tasks they need to perform. Some examples of these situations are discussed below.

In 1998 the pilots of a Boeing 757 failed to notice that the auto-throttle system had disengaged. The pilots sensed a slight vibration, and after detecting a dangerously low airspeed, the captain correctly attributed the vibration to a loss of lift by the wings. To regain the required airspeed, the throttles were advanced and a slow descent was initiated. However, upon descent the aircraft nearly collided with another plane and both planes needed to be instructed to adopt new courses. The captain claimed that no warning had been provided to alert the crew that the automatic throttle system had disengaged.

In the aftermath of the crash of the American Airlines flight 965 near Cali, Columbia, in 1995, the FAA's human factors team suggested that pilots might not know how to interpret computer system information. The pilots of that flight accepted an offer to land on a different runway, forcing them to rush their descent. In the process, they incorrectly programmed their FMS to direct their plane to Bogota, which was off course by more than 30 miles, and ultimately flew into a 9000-foot mountain.

On a normal approach into Nagoya, Japan, in 1994, the first officer of a China Airlines Airbus A-300 hit the wrong switch on the autopilot, sending the plane into an emergency climb. The throttles increased automatically and the nose pitched up. As the pilots reduced power and tried to push the nose down, the flightdeck computers became even more determined to make the plane climb. The nose rose to 53 degrees, and despite adding full power, the airspeed dropped to 90 mph, which was too slow to maintain the plane in the air. The aircraft crashed tail first into the ground near the runway.

In 1998 a Boeing 737 bound for Denver was instructed by air traffic controllers to descend quickly to 19,000 feet to avoid an oncoming plane. The captain attempted to use the FMS to execute the descent, but the system did not respond quickly enough. Following a second order by air traffic controllers, the captain opted to turn off the FMS and assume manual control, resulting in a near miss with the other plane. The captain attributed his "error" to reliance on automation.

7.3 Adapting to Automation and New Technology

7.3.1 Designer Error

As is the case with user performance of various types of products, the performance of designers will also depend on the operational contexts in which they are working and will be susceptible to many of the same forms of errors (Smith and Geddes, 2003). Working against designers is the increased specialization and heterogeneity of work domains, which is making it exceedingly difficult for them to anticipate the effects on users of introducing automation and new technologies. Nonetheless, errors resulting from user interactions with new technologies are now often attributed to designers. Designer errors could arise from inadequate or incorrect knowledge about the application area (i.e., a failure for designers to anticipate important scenarios) or the inability to anticipate how the product will influence user performance (i.e., insufficient understanding by designers).

In reality, designers' conceptualizations are nothing more than initial hypotheses concerning the

*This section is adapted from Kaye (1999a).

collaborative relationship between their technological product and the human. Accordingly, their beliefs regarding this relationship need to be gradually shaped by data that are based on actual human interaction with these technologies, including the transformations in work experiences that these interactions produce (Dekker, 2005). However, as Dekker notes, in practice the validation and verification studies by designers are usually limited, providing results that may be informative but "hardly about the processes of transformation (different work, new cognitive and coordination demands) and adaptation (novel work strategies, tailoring of the technology) that will determine the sources of a system's success and potential for failure once it has been fielded" (p. 164). In the study on computerized physician order-entry systems discussed in Section 3.4, many of the errors that were identified were probably rooted in constraints of these kinds that were imposed on the design process.

Although designers have a reasonable number of choices available to them that can translate into different technical, social, and emotional experiences for users, like users they themselves are under the influence of sociocultural (Evan and Manion, 2002) and organizational factors (Figure 1). For example, the reward structure of the organization, an emphasis on rapid completion of projects, and the insulation of designers from the consequences of their design decisions can induce designers to give less consideration to factors related to ease of operation and even safety (Perrow, 1983). Although these circumstances would appear to shift the attribution of user errors from designers to management, designer errors and management errors both represent types of *latent errors* that are responsible for creating the preconditions for user errors (Reason, 1990). Perrow (1999) contends that a major deficiency in the design process is the inability of designers and management to appreciate human fallibility by failing to take into account relevant information that could be supplied by human factors and ergonomics specialists. This concern is given serious consideration in user-centered design practices (Nielsen, 1995). However, in some highly technical systems where designers may still be viewing their products as closed systems governed by perfect logic, this issue remains unresolved. The way the FAA has approached this problem has been through recommendations to manufacturers that they make displays and controls easier to use and that they develop a better understanding of pilot vulnerabilities to complex environments. For example, in Boeing's modern air fleets, all controls and throttles provide visual and tactile feedback to pilots—thus the control column that a pilot normally pulls back to initiate climb will move back on its own when a plane is climbing on autopilot.

7.3.2 The Keyhole Property, Task Tailoring, and System Tailoring

Much of our core human factors knowledge concerning human adaptation to new technology in complex systems is derived from experiences in the nuclear power and aviation industries. These industries were forced to address the consequences of imposing on their workers major transformations in the way that system data were presented. In nuclear power control rooms, the banks of hardwired displays were replaced by one or a few computer-based display screens, and in cockpits the analog single-function single displays were replaced by sophisticated software-driven electronic integrated displays. These changes drastically altered the human's visual–spatial landscape and offered a wide variety of schemes for representing, integrating, and customizing data. For those experienced operators who were used to having the entire data world available to them at a glance, adapting to the new technology was far from straightforward. The mental models and strategies that were developed based on having all system state information available simultaneously were not likely to be as successful when applied to these newly designed environments, making these operators more predisposed to errors than were their less experienced counterparts.

In complex work domains such as health care that require the human to cope with a potentially enormous number of different task contexts, anticipating the user's adaptation to new technology can become so difficult for designers that they themselves, like the practitioners who will use their products, can be expected to conform to strategies of minimizing cognitive effort. Instead of designing systems with operational contexts in mind, a cognitively less taxing solution is to identify and make available all possible information that the user may require but to place the burden on the user to search for, extract, or configure the information as the situation demands. These designer strategies are often manifest as computer mediums that exhibit the *keyhole property*, whereby the size of the available viewports (e.g., windows) is very small relative to the number of data displays that potentially could be examined (Woods and Watts, 1997). Unfortunately, this approach to design makes it more likely that the user can "get lost in the large space of possibilities" and makes it difficult to find the right data at the right time as activities change and unfold.

In a study by Cook and Woods (1996) on adapting to new technology in the domain of cardiac anesthesia, physiological monitoring equipment dedicated to cardiothoracic surgery was upgraded from separate devices to a computer system that integrated the functions of four devices onto a single color display. However, the flexibilities that the new technology provided in display options and display customization also created the need for physicians to direct attention to interacting with the patient monitoring system. By virtue of the keyhole property there were now new interface management tasks to contend with. These tasks derived in part from the need to access highly interrelated data serially, thus potentially degrading the accuracy and efficiency of the mental models the physicians required for making patient intervention decisions. New interface management tasks also included the need to declutter displays periodically to avoid obscuring data channels that required monitoring. This requirement resulted from collapsing into a single device the data world previously made available by the multi-instrument configuration.

To cope with these potentially overloading situations, physicians were observed to tailor both the computer-based system (*system tailoring*) and their own cognitive strategies (*task tailoring*). For example, the physicians discovered that the default blood pressure display configuration for the three blood pressures that were routinely displayed was unsuitable—the waveforms and numeric values (derived from digital processing) changed too slowly and eliminated important quantitative information. Rather than exploit the system's flexibility, the physicians simplified the system by constraining the display of data into a fixed spatially dedicated default organization. This required substantial effort, initially to force the preferred display configuration prior to the initiation of a case, then to ensure that this configuration is maintained in the event that the computer system performs automatic window management functions. To tailor their tasks, they planned their interactions with the device to coincide with self-paced periods of low criticality, and developed stereotypical routines to avoid getting lost in the complex menu structures rather than exploiting the system's flexibility. In the face of circumstances incompatible with task-tailoring strategies, which are bound to occur in this complex work domain, the physicians had no choice but to confront the complexity of the device, thus diverting information-processing resources from the patient management function (Cook and Woods, 1996). This irony of automation, whereby the burden of interacting with the technology tends to occur during those situations when the human can least afford to divert attentional resources, is also found in aviation. As noted, automation in cockpits can potentially reduce workload by allowing complete flight paths to be programmed through keyboards. Changes in the flight path, however, require that pilots divert their attention to the numerous keystrokes that need to be input to the keyboard, and these changes tend to occur during takeoff or descent—the phases of flight containing the highest risk and that can least accommodate increases in pilot workload (Strauch, 2002).

Task tailoring reflects a fundamental human adaptive process. Thus, humans should be expected to shape new technology to bridge gaps in their knowledge of the technology and fulfill task demands. The concern with task tailoring is that it can create new cognitive burdens, especially when the human is most vulnerable to demands on attention, and mask the real effects of technology change in terms of its capability for providing new opportunities for human error (Dekker, 2005).

7.3.3 Effects of New Technology on Team Communication

Cao and Taylor (2004) recently examined the effects of introducing a remote surgical robot on communication among the operating room (OR) team members. Understanding the potential for human errors brought about from interactions among team members in the face of this new technology requires closely examining contextual factors such as communication, teamwork, flow of information, work culture, uncertainty, and overload (Figure 1). In their study, a framework referred to as *common ground* (Clark and Schaefer, 1989) was used to analyze communication for two cholecystectomy procedures that were performed by the same surgeon: one using conventional laparoscopic instruments and the other using a robotic surgical system. Common ground represents a person's knowledge or assumptions about what other people in the communication setting know. It can be established through past and present experiences in communicating with particular individuals, the knowledge or assumptions one has about those individuals, and general background information. High levels of common ground would thus be expected to result in more efficient and accurate communication.

In the OR theater, common ground can become influenced by a number of factors. For instance, the surgeon's expectations for responses by team members may depend on the roles (such as nurse, technician, or anesthesiologist) that those persons play. Other factors that can affect the level of common ground include familiarity with team members, which is often undermined in the OR due to rotation of surgical teams, and familiarity with the procedure. When new technology is introduced, all these factors conspire to erode common ground and thus potentially compromise patient safety. Roles may change, people become less familiar with their roles, the procedures for using the new technology are less familiar, and expectations for responses from communication partners becomes more uncertain. Misunderstandings can propagate through team members in unpredictable ways, ultimately leading to new forms of errors.

The introduction of a remote master–slave surgical robot into the OR necessitates a physical barrier, and what Cao and Taylor (2004) observed was that the surgeon, now removed from the surgical site, had to rely almost exclusively on video images from this remote surgical site. Instead of receiving a full range of sensory information from the visual, auditory, haptic, and olfactory senses, the surgeon had to contend with a "restricted field of view and limited depth information from a frequently poor vantage point" (p. 310) and increased uncertainty regarding the status of the remote system. These changes potentially overload the surgeon's visual system and create more opportunities for decision-making errors, due to gaps in the information that is being received. Also, in addition to the need for obtaining information on patient status and the progress of the procedure, the surgeon had to cope with information-processing demands deriving from the need to access information about the status of the robotic manipulator. Thus, to ensure effective coordination of the procedure, the surgeon was now responsible for verbally distributing more information to the OR team members than with conventional laparoscopic surgery.

Overall, significantly more communication within the OR team was observed under robotic surgery conditions than with conventional laparoscopic surgery. Moreover, the communication patterns were haphazard, which increased the team member's uncertainty concerning when information and what information should be distributed or requested and thereby the potential for human error resulting from miscommunication and lack of communication. Use of different terminologies in referring to the robotic system and startup confusion contributed to the lack of common ground. Although training on the use of this technology was provided to these surgical team members, the findings suggested the need for training to attain common ground. This could possibly be achieved through the use of rules or an *information visualization system* that could facilitate the development of a shared mental model among the team members (Stout et al., 1999).

8 HUMAN ERROR IN MAINTENANCE ACTIVITIES

To function effectively, almost all systems require maintenance. Most organizations require both scheduled (preventive) maintenance and unscheduled (active) maintenance. Whereas unscheduled maintenance is required when systems or components fail, preventive maintenance attempts to anticipate failures and thereby minimize system unavailability. Frequent scheduled maintenance can be costly, and organizations often seek to balance these costs against the risks of equipment failures. Lost in this equation, however, is a possible "irony of maintenance"—that an increased frequency in scheduled maintenance may actually increase system risk by providing more opportunities for human interaction with the system (Reason, 1997). This increase in risk is more likely if assembly rather than disassembly operations are called for, as the comparatively fewer constraints associated with assembly operations makes these activities much more susceptible to various errors, such as identifying the wrong component, applying inappropriate force, or omitting an assembly step.

Maintenance environments are notorious for breakdowns in communication, often in the form of implicit assumptions or ambiguity in instructions that go unconfirmed (Reason and Hobbs, 2003). When operations extend over shifts and involve unfamiliar people, these breakdowns in communication can propagate into catastrophic accidents, as was the case in the explosion aboard the Piper Alpha oil and gas platform in the North Sea (Reason and Hobbs, 2003) and the crash of ValueJet flight 592 (Strauch, 2002). Incoming shift workers are particularly vulnerable to errors following commencement of their task activities, especially if maintenance personnel in the outgoing shift conclude their work at an untimely point in the procedure and fail to brief incoming shift workers adequately as to the operational context about to be confronted (Sharit, 1998). In these cases, incoming shift workers are placed in the difficult position of needing to invest considerable attentional resources almost immediately in order to avoid an incident or accident.

Many preventive maintenance activities initially involve searching for flaws prior to applying corrective procedures, and these search processes are often subject to various expectancies that could lead to errors. For example, if faults or flaws are seldom encountered, the likelihood of missing such targets will increase; if they are encountered frequently, properly functioning equipment may be disassembled. Maintenance workers are also often required to work in restricted spaces that are error inducing by virtue of the physical and cognitive constraints that these work conditions impose (Reynolds-Mozrall et al., 2000).

Flawed partnerships between maintenance workers and troubleshooting equipment can also give rise to errors. As with other types of automation or aiding devices, troubleshooting aids can compensate for human limitations and extend human capabilities when designed appropriately. However, these devices are often opaque and may be misused or disregarded (Parasuraman and Riley, 1997), depending on the worker's self-confidence, prior experiences with the aid, and knowledge of co-worker attitudes toward the device. For instance, if the logic underlying the software of an expert troubleshooting system is inaccessible, the user may not trust the recommendations or explanations given by the device (Section 7.1) and therefore choose not to replace a component that the device has identified as faulty.

Errors resulting from interruptions are particularly prevalent in maintenance environments. Interruptions due to the need to assist a co-worker or following the discovery that the work procedure called for the wrong tool or equipment generally require the worker to leave the scene of operations, and the most likely error in these types of situations is an omission. In fact, memory lapses probably constitute the most common errors in maintenance, suggesting the need for incorporating good reminders (Table 12). Reason and Hobbs (2003) emphasize the need for mental readiness and mental rehearsal as ways that maintenance workers can inoculate themselves against errors that could arise from interruptions, time pressure, communication, and unfamiliar situations that may arise.

Written work procedures are pervasive in maintenance operations, and there may be numerous problems with the design of these procedures that can predispose their users to errors (Drury, 1998). Violations of these procedures are also relatively common, and management has been known to consider such violations as causes and contributors of adverse events—a belief that is both simplistic and unrealistic. The assumptions that go into the design of procedures are typically based on normative models of work operations. However, the actual contexts under which real work takes place are often very different from those that the designers of the procedures have envisioned or were willing to acknowledge. To the followers of the procedures, who must negotiate

Table 12 Characteristics of Good Reminders

Universal Criteria

- *Conspicuous.* It should be able to attract the person's attention at the critical time.
- *Contiguous.* It should be located as closely as possible in both time and distance to the to-be-remembered (TBR) task step.
- *Context.* It should provide sufficient information about when and where the TBR step should be carried out.
- *Content.* It should inform the person about what has to be done.
- *Check.* It should allow the person to check off the number of discrete actions or items that should be included in correct performance of the task.

Secondary Criteria

- *Comprehensive.* It should work effectively for a wide range of TBR steps.
- *Compel.* It should (when warranted or possible) block further progress until a necessary prior step has been completed.
- *Confirm.* It should help the person to establish that the necessary steps have been completed. In other words, it should continue to exist and be visible for some time after the performance of the step has passed.
- *Conclude.* It should be readily removable once the time for the action and its checking have passed.

Source: Adapted from Reason (1997).

their tasks while being subjected to limited resources, conflicting goals, and pressures from various sources, the cognitive process of transforming procedures into actions is likely to expose incomplete and ambiguous specifications that at best appear only loosely related to the actual circumstances (Dekker, 2005). A worker's ability to adapt (and thereby violate) these procedures successfully may in fact be lauded by management and garner respect from fellow workers. However, if these violations happen to become linked to accidents, management would probably deny the existence of any unspoken approval of these informal activities and retreat to the official doctrine: Safety can result only if workers follow procedures.

As indicated by Dekker (2005), skilled workers who attempt to adapt procedures to the situation face a double bind: "If rote rule following persists in the face of cues that suggest procedures should be adapted, this may lead to unsafe outcomes. People can get blamed for their inflexibility, their application of rules without sensitivity to context. If adaptations to unanticipated conditions are attempted without complete knowledge of circumstance or certainty of outcome, unsafe results may occur too. In this case, people get blamed for their deviations, their nonadherence" (p. 140). Dekker suggests that organizations monitor (Section 6.5) and understand the basis for the gaps between procedures and practice and develop ways of supporting the cognitive skill of applying procedures successfully across different situations by enhancing workers' judgments of when and how to adapt.

9 ORGANIZATIONAL AND WORK GROUP CULTURES

As with people who live in the same regions or share similar religious beliefs, members of groups within companies, such as maintenance workers, control room operators, or workers involved in transporting goods can also embody beliefs and practices that reflect their shared values. These various work group cultures can be influenced by select individuals who choose to impose their views on subordinates, as well as by the norms that characterize the entire organization. Although cultural factors associated with the organization are generally assumed to be responsible for the norms adopted by work group cultures, in reality organizational culture can have varying degrees of influence on the development and behavior of any particular work group culture.

Strauch (2002) has noted that cultural factors "can make the difference between effective and erroneous performance" (p. 111), and identified two cultural antecedents to error: acceptance of authority and identification with the group. In Hofstede's (1991) analysis of the influence of company cultures on behaviors among individuals, identification with the group was termed *individualism–collectivism*, and acceptance of authority was referred to as *power distance*. Whereas individually oriented people place personal goals ahead of organizational goals, collectivist-oriented persons tend to identify with the company (or work group), so more of the responsibility for errors that they commit would be deflected onto the company. These distinctions thus underlie attitudes that can possibly affect the degree to which workers prepare mentally for potential errors (Section 11).

Power distance refers to the differences in power that employees perceive between themselves and subordinates and superiors. In cultures with high power distance, subordinates are less likely to point out or comment to others about errors committed by superiors as compared to workers in company cultures with low power distance. Although differences in power distance tend to be associated with different countries, this factor can have a considerable impact in ethnically diverse organizations that have become commonplace in many Western societies. Thus, in a Canadian hospital a nurse originating from and trained in the Philippines, a country with a relatively high power distance score, may be less willing than her Canadian counterpart to question a possibly incorrect medication order by a physician. Cultures in which workers tend to defer to authority can also suppress the organization's capability for learning. For example, workers may be less willing to make suggestions that can improve training programs or operational procedures (Section 6.2).

A third cultural factor identified by Hofstede, *uncertainty avoidance*, refers to the willingness or ability to deal with uncertainty. This factor also has implications for human error. For example, workers in cultures that are low in uncertainty avoidance are probably more likely to invoke performance at the knowledge-based level (Section 3.2.5) in response to novel or unanticipated situations for which rules are not available.

Can companies with good cultures be differentiated from those with bad cultures? High-reliability organizations (Section 11) that anticipate errors and encourage safety at the expense of production, that have effective error-reporting mechanisms without fear of reprisals, and that maintain channels of communication across all levels of the company's operations generally reflect good cultures. Questionable hiring practices, poor economic incentives, inflexible and outmoded training programs, the absence of incident reporting systems and meaningful accident investigation mechanisms, managerial instability, and the promotion of atmospheres that discourage communication between superiors and subordinates are likely to produce poor organizational and work group cultures.

Errors associated with maintenance operations can often be traced to organizational culture. This was clearly the case in the crash of ValueJet flight 592 into the Florida Everglades in 1996 just minutes after takeoff. The crash occurred following an intense fire in the airplane's cargo compartment that made its way into the cabin and overcame the crew (Strauch, 2002). Unexpended and unprotected canisters of oxygen generators, which can inadvertently generate oxygen and heat and consequently ignite adjacent materials, had somehow managed to become placed onto the aircraft. Although most of the errors that were uncovered by the investigation were associated with maintenance technicians at SabreTech—the maintenance facility contracted by ValueJet to overhaul several of its aircraft—these errors were attributed to practices at SabreTech that reflected organizational failures. Specifically, the absence of information on the work cards concerning the removal of oxygen canisters from two ValueJet airplanes that were being overhauled led to the failure by maintenance personnel to lock or expend the generators. There was also a lack of communication across shifts concerning the hazards associated with the oxygen generators; although some technicians who had removed the canisters from the other aircraft knew of the hazards, others did not. In addition, procedures for briefing incoming and outgoing shift workers concerning hazardous materials and for tracking tasks performed during shifts were not in place. Finally, parts needed to secure the generators were unavailable, and none of the workers in shipping and receiving, who were ultimately responsible for placing the canisters on the airplane, was aware of the hazards.

Relevant to this discussion was the finding that the majority of the technicians that removed oxygen canisters from ValueJet airplanes as part of the overhaul of these aircraft were not SabreTech personnel but contractor personnel. In the absence of an adequately informed organizational culture, it comes as no surprise that management would be oblivious to the implications of outsourcing on worker communication and task performance. Further arguments concerning the importance of organizational culture for system safety can be found in Reason (1997) and Vicente (2004).

9.1 The *Columbia* Accident

The *Columbia* space shuttle accident in 2003 exposed a failed organizational culture. The physical cause of the accident was a breach in the thermal protection system on the leading edge of *Columbia*'s left wing about 82 seconds after the launch. This breach was caused by a piece of insulating foam that separated from the external tank in an area where the orbiter attaches to the external tank. The *Columbia* Accident Investigation Board's (2003) report stated that "NASA's organizational culture had as much to do with this accident as foam did," that "only significant structural changes to NASA's organizational culture will enable it to succeed," and that NASA's current organization "has not demonstrated the characteristics of a learning organization" (p. 12).

To some extent NASA's culture was shaped by compromises with political administrations that were required to gain approval for the space shuttle program. These compromises imposed competing budgetary and mission requirements that resulted in a "remarkably capable and resilient vehicle" but one that was "less than optimal for manned flights" and "that never met any of its original requirements for reliability, cost, ease of turnaround, maintainability, or, regrettably, safety" (p. 11). The organizational failures are almost too numerous to document: unwillingness to trade off scheduling and production pressures for safety; shifting management systems and a lack of integrated management across program elements; reliance on past success as a basis for engineering practice rather than on dependable engineering data and rigorous testing; the existence of organizational barriers that compromised communication of critical safety information and discouraged differences of opinion; and the emergence of an informal command and decision-making apparatus that operated outside the organization's norms. According to the *Columbia* Accident Investigation Board, deficiencies in communication, both up and down the shuttle program's hierarchy, were a foundation for the *Columbia* accident.

These failures were largely responsible for missed opportunities, blocked or ineffective communication, and flawed analysis by management during *Columbia*'s final flight that hindered the possibility of a challenging but conceivable rescue of the crew by launching the *Atlantis*, another space shuttle craft, to rendezvous with *Columbia*. The accident investigation board concluded: "Some Space Shuttle Program managers failed to fulfill the implicit contract to do whatever is possible to ensure the safety of the crew. In fact, their management techniques unknowingly imposed barriers that kept at bay both engineering concerns and dissenting views, and ultimately helped create 'blind spots' that prevented them from seeing the danger the foam strike posed" (p. 170). Essentially, the position adopted by managers concerning whether the debris strike created a safety-of-flight issue placed the burden on engineers to prove that the system was unsafe.

Numerous deficiencies were also found with the Problem Reporting and Corrective Action database, a critical information system that provided data on any nonconformances. In addition to being too time consuming and cumbersome, it was also incomplete. For example, only foam strikes that were considered in-flight anomalies were added to this database, which masked the extent of this problem.

Finally, what is particularly disturbing was the failure of the shuttle program to detect the foam trend and appreciate the danger that it presented. Shuttle managers discarded warning signs from previous foam strikes and normalized their occurrences. In so doing, they desensitized the program to the dangers of foam strikes and compromised the flight readiness process. Many workers at NASA knew of the problem. However, in the absence of an effective mechanism for communicating these "incidents" (Section 6) proactive approaches for identifying and mitigating risks were unlikely to be in place. In particular, a proactive perspective to risk identification and management could have resulted in a better understanding of the risk of thermal protection damage from foam strikes, tests being performed on the resilience of the reinforced carbon–carbon panels, and either the elimination of external tank foam loss or its mitigation through the use of redundant layers of protection.

10 INVESTIGATING HUMAN ERROR IN ACCIDENTS AND INCIDENTS

10.1 Causality and Hindsight Bias

Investigations of human error are generally performed as part of accident and incident investigations (Chapter 41). In conducting these investigations, the most fundamental issue is the attribution of causality to incident and accident events. Currently, there are a variety of techniques that investigators can choose from to assist them in performing causal analysis (Johnson, 2003).

A related issue concerns the level of detail required for establishing causality (Senders and Moray, 1991). At one level of analysis a cause can be an interruption; at a different level of analysis the cause can be attributed to a set of competing neural activation patterns that result in an action slip. Dilemmas regarding the appropriate level of causal analysis are usually resolved by considering the requirements of the investigative analysis. Generally, analysts can be expected to employ heuristics such as satisficing (Section 3.2), whereby decisions and judgments are made that appear good enough for the purposes of the investigation. Investigators also need to be aware of the possible cognitive biases that reporters of incidents and accidents may be harboring (Table 11). These same biases can also play a part in how witnesses attribute blame and thus in how they perceive relationships between causes and effects.

What makes determining causes of accidents especially problematic for investigators is that they typically work with discrete fragments of information derived from decomposing continuous and interacting sequences of events. This ultimately leads to various distortions in the true occurrence of events (Woods, 1993). A further complication is *hindsight bias*, which derives from the tendency to judge the quality of a process based on whether positive or negative outcomes ensued (Fischhoff, 1975; Christoffersen and Woods, 1999). Because accident investigators usually have knowledge about negative outcomes, through hindsight they can look back and identify all the failed behaviors that are consistent with these outcomes (Section 1.1). It is then highly probable that a causal sequence offering a crisp explanation for the incident will unfold to the investigator. A foray through Casey's (1993) reconstructed accounts of a number of high-profile accidents attributed to human error would probably transform many in the lay public into hindsight experts.

Although hindsight bias can assume a number of forms (Dekker, 2001), they all derive from the tendency to treat actions in isolation and thus distort the context in which the actions took place. The pervasiveness of the hindsight bias has led Dekker (2005) to suggest the intriguing possibility that it may actually be serving an adaptive function—that is, the hindsight bias is not so much about explaining what happened as it is about future survival, which would necessarily require the decontextualization of past failures into a "linear series of binary choices." The true bias thus derives from the belief that the oversimplification of rich contexts into a series of clearly defined choices will increase the likelihood of coping with complexity successfully in the future. However, in reality, by obstructing efforts at establishing cause and effect, the hindsight bias actually jeopardizes the ability to learn from accidents (Woods et al., 1994) and thus the ability to predict or prevent future failures.

10.2 Methods for Investigating Human Error

Investigations of human error can be pursued using either informal approaches or methods that are much more systematic or specialized. Strauch's (2002) approach, which reflects a relatively broad and informal perspective to this problem, emphasizes antecedent factors (e.g., equipment, operator, maintenance, and cultural factors), data collection and analysis issues, and factors such as situation awareness and automation, all of which are interwoven with case studies. The more specialized methods for investigating human error are generally referred to as accident or incident analysis techniques, although in some cases these tools can also be used to assess the potential risks associated with systems or work processes. Examples of techniques used exclusively for investigating accidents include change analysis (Kepner and Tregoe, 1981) and the sequentially timed events plotting procedure (Hendrick and Benner, 1987).

Change analysis techniques are based on the well-documented general relationship between change and increased risk. These techniques make use of accident-free reference bases to identify systematically changes or differences associated with the accident or incident situation. A simple worksheet is usually all that is required for exploring potential changes contributory to adverse outcomes. Listed in the rows of the first column of this worksheet are factors that are stated in terms of questions regarding who, where, what, when, how, and why with respect to task factors, working conditions, initiating events, and management control factors. The next columns, respectively, address each of these factors in terms of the present (accident or incident) situation, prior (accident-free) situation, a comparison of these two situations in order to identify changes or differences, and a list of the differences. Finally, differences are analyzed for their effect on the accident or incident in terms of both their independent and interactive contributions. When used in conjunction with TA or CTA methods, change analysis can be applied both retrospectively to facilitate the identification of underlying causes of human error, and proactively to predict adverse consequences by investigating potential problems associated with proposed changes in normal or stable functioning systems.

Another well-known technique, the management oversight and risk tree (MORT), relies on a logic diagram for investigating the various factors contributing to an accident (Johnson, 1980). Factors considered by MORT include lines of responsibility, barriers to unwanted energy sources, and management factors. By reasoning backward through a sequence of contributory factors, responding "yes" and "no" to questions along the way, and through the availability of accompanying text that aids the analyst in judging whether a factor is adequate or less than adequate, MORT assists the analyst in detecting omissions, oversights, or defects, and may be especially useful for identifying organizational root causes (Gertman and Blackman, 1994).

The accident investigation method that has recently been given the most attention is *root-cause analysis* (RCA). This method usually refers to the formal application of a root cause decision tree diagram for investigating why a particular event occurred. In the SOURCE (seeking out the underlying root causes of events) method of performing RCA (ABS Consulting Group, 1998), *root causes* are defined as "the most basic causes that can reasonably be identified, which management has control to fix and for which effective recommendations for preventing recurrence can be generated" (p. 2). The major steps in the SOURCE RCA process are illustrated in Figure 11. The first step, data collection, represents the most time-consuming step of the process. Although data collection is assumed to occur throughout the analysis, data from relatively unstable sources, such as people and certain types of physical data, need to be collected as soon as possible. The interviewing technique employed by the investigator will probably be the most critical factor in determining the effectiveness of data gathering (Strauch, 2002).

The next step, *causal factor charting*, utilizes a causal factor chart to describe in sequence the events leading up to and following the incident, as well as the conditions surrounding these events. A skeletal causal chart is generated based on the initial data collected; this chart is then modified progressively as data accumulate. Other elements in addition to those illustrated in Figure 11 can be incorporated into causal factor charts, including presumptive events, presumptive conditions, presumptive causal factors, and items of note. A number of principles, guidelines, and procedures are offered for supporting the causal factor charting process (ABS Consulting Group, 1998).

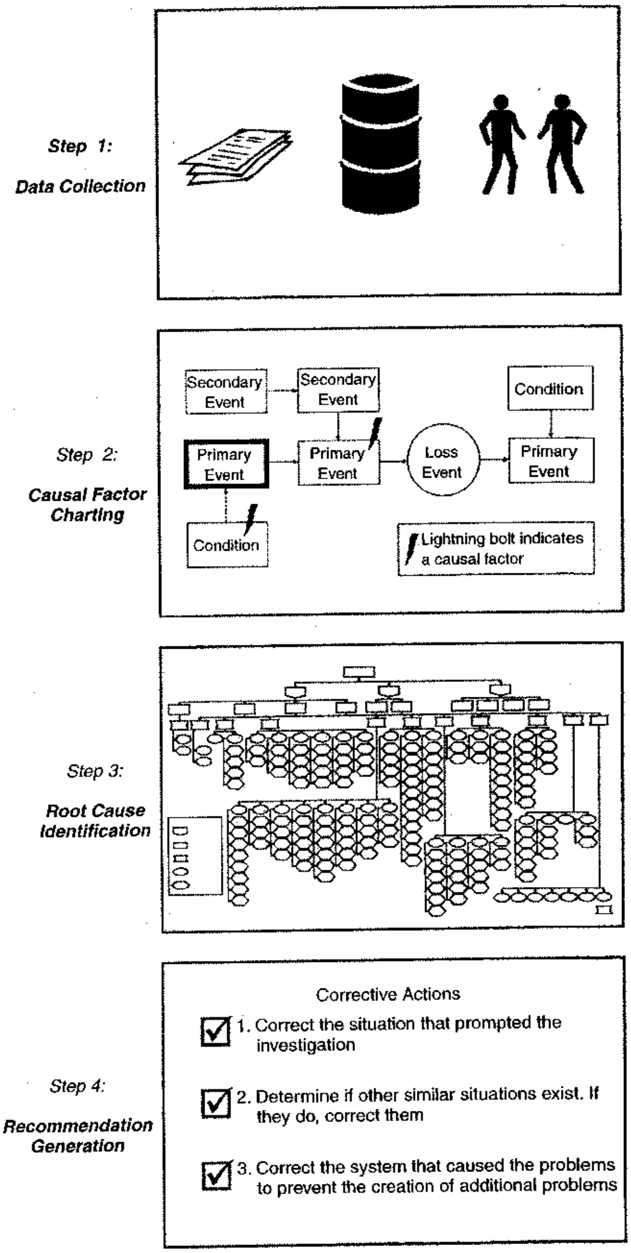

Step 1:
Data Collection
Step 2:
Causal Factor Charting
Secondary Event
Secondary Event
Condition
Primary Event
Primary Event
Loss Event
Primary Event
Condition
Lightning bolt indicates a causal factor
Step 3:
Root Cause Identification
Step 4:
Recommendation Generation
Corrective Actions
☑ 1. Correct the situation that prompted the investigation
☑ 2. Determine if other similar situations exist. If they do, correct them
☑ 3. Correct the system that caused the problems to prevent the creation of additional problems

Figure 11 Root-cause analysis method. (From ABS Consulting Group, 1998.)

The third step of this process involves the use of a tree diagram called a *root-cause map* to identify the underlying reasons for each causal factor identified during causal factor charting. For each causal factor the analyst must determine which top-level node in the map is most applicable. Based on this decision, the analyst then works down through the lower (more specific) levels of the map, selecting the most applicable node at each level. The three upper-level nodes of the map correspond to equipment failures, personnel failures, and other failures; however, only the first two categories are analyzed for root causes. At the second level these three nodes are subdivided into 10 problem category nodes. Examples of these categories are equipment design problem, equipment misuse, contract employee, natural phenomena, and sabotage or horseplay. The third level of the map consists of nodes corresponding to 11 major root-cause categories; examples of these categories include procedures, human factors engineering, training, and communications. In the transition from the second to the third level, the map allows for a number of points of intersection between equipment failures and personnel failures, thus allowing all failures to be traced back to some type of human error. At the fourth level of the map these categories become subdivided into near root causes, which in turn are subdivided at the bottom level into a detailed set of root causes. To aid the investigator in using the root-cause map, examples for each node are provided in terms of typical issues and typical recommendations.

At this point in the process, a root-cause summary table can be generated that links each causal factor in the chart with one or more paths in the map that terminate at root causes and to recommendations that address each of these root causes. These tables then form the basis for investigative reports that comprise the final step of RCA.

The availability of a systematic method for performing incident and accident investigations within a high-risk organization will increase an organization's potential for learning, improvement, and development of positive work cultures. However, as with IRSs these benefits are anticipated only when these investigations are not used as a basis for reprisals and when workers are informed about and involved in the investigative process.

11 TOWARD MINIMIZATION OF HUMAN ERROR AND THE CREATION OF SAFE SYSTEMS

Human error is a complex phenomenon. Recent evidence from neuroimaging studies has linked an *error negativity*, an event-related brain potential probably originating from the anterior cingulate cortex, to the detection by individuals of action slips, errors of choice, and other errors (Nieuwenhuis et al., 2001; Holroyd and Coles, 2002), possibly signifying the existence of a neurophysiological basis for a preconscious action-monitoring system. However, suggestions that these kinds of findings may offer possibilities for predicting human errors in real-time operations (Parasuraman, 2003) are probably overstated. Event-related brain potentials may provide insight into attentional preparedness and awareness of response conflicts, but the complex interplay of factors responsible for human error (Section 3.1) takes these discoveries out of contention as meaningful explanatory devices.

Although managers often speak in terms of the need for eliminating human error, this goal is neither desirable nor reasonable. The benefits that derive from the realization that errors have been committed should not be readily dismissed; they play a critical role in human adaptability, creativity, and the manifestation of expertise. The elimination of human error is also inconceivable if only because human fallibility will always exist. Tampering with human fallibility, for example by increasing the capabilities of working memory and attention, would probably facilitate the design and production of new and more complex systems, and ultimately, new and unanticipated opportunities for human error. More realistically, the natural evolution of knowledge and society should translate into the emergence of new systems, new forms of interaction among people and devices, and new sociopolitical and organizational cultures that will, in turn, provide new opportunities for enabling human fallibility.

However, in no way should these suppositions detract from the goal of human error reduction, especially in complex high-risk systems. As a start, system hardware and software need to be made more reliable, better partnerships between humans and automation need to be established, barriers that are effective in providing detection and absorption of errors without adversely affecting contextual and cognitive constraints need to be put in place, and incident-reporting systems that enable organizations to learn and anticipate, especially when errors become less frequent and thus deprive analysts with the opportunity for preparing and coping with their effects, need to become more ubiquitous.

Organizations also need to consider the impact that various economic incentives may have on shaping work behaviors (Moray, 2000) and the adoption of strategies and processes for implementing features that have come to be associated with *high-reliability organizations* (HROs) (Rochlin et al., 1987; Roberts, 1990). By incorporating fundamental characteristics of HROs, particularly the development of cultures of reliability that anticipate and plan for unexpected events, try to monitor and understand the gap between work procedures and practice (Dekker, 2005), and place value in organizational learning, the adverse consequences of interactive complexity and tight coupling that Perrow's theory predicts (Section 3.3) can largely be countered.

In addition, methods for describing work contexts and for determining and assessing the perceptions and assessments that workers make in response to these contexts, as well as rigorous TA and CTA techniques for determining the possible ways that fallible humans can become ensnared by these situations, need to be investigated, implemented, and continuously evaluated in order to strengthen the predictive capabilities of human error models. These methods also need to be integrated into the conceptual, development, and testing stages of the design process to better inform designers (of both products and work procedures) about the potential effects of design decisions, thus bridging the gap between the knowledge and intentions of the designer and the needs and goals of the user.

Problems created by poor designs and management policies traditionally have been dumped on training departments (CCPS, 1994). Instead of using training to compensate for these problems, it should be given a proactive role in minimizing, detecting, and recovering errors. This can be achieved through innovative training methods that emphasize management of task activities under uncertainty and time constraints; integrate user-centered design principles for establishing performance support needs (such as the need for planning aids); give consideration to the kinds of cues that are necessary for developing situation awareness (Endsley et al., 2003) and for interpreting common-cause and common-mode system failures; and utilize simulation methods effectively for providing extensive exposure to a wide variety of contexts. By including provisions in training for imparting mental preparedness, people will be better able to anticipate the anomalies they might encounter and the errors they might make, and to develop error-detection skills (Reason and Hobbs, 2003).

Although worker selection (Chapter 17) is a potentially explosive issue, it can be used to exploit individual variability in behavioral tendencies and cognitive capabilities and thus provide better human–system fits (Damos, 1995). Bierly and Spender (1995) have documented the extraordinary safety record of the U.S nuclear navy and attributed it in part to a culture that insisted on careful selection of people who were highly intelligent, very motivated, and who were then thoroughly trained and held personally accountable for their tasks. These characteristics created the work culture context for communications that could: be carried out under conditions of high risk and high stress; flow rapidly either top-down or bottom-up through the chain of command; and encompass information about mistakes, whether technical, operational, or administrative, without fear of reprisals.

However, perhaps the greatest challenge in reducing human error is managing these error-management processes (Reason and Hobbs, 2003)—defense strategies need to be aggregated coherently (Amalberti, 2001). Too often these types of error-reduction enterprises, innovative as they may be, remain isolated or hidden from each other. This needs to change—all programs that can influence error management need to be managed as a unified synergistic entity.

REFERENCES

ABS Consulting Group (1998), *Root Cause Analysis Handbook: A Guide to Effective Incident Investigation*, Government Institutes Division, Knoxville, TN.

Amalberti, R. (2001), "The Paradoxes of Almost Totally Safe Transportation System," *Safety Science*, Vol. 3, pp. 109–126.

Bainbridge, L. (1987), "Ironies of Automation," in *New Technology and Human Error*, J. Rasmussen, K. Duncan, and J. Leplat, Eds., Wiley, New York, pp. 273–276.

Barach, P., and Small, S. (2000), "Reporting and Preventing Medical Mishaps: Lessons from Non-medical Near Miss Reporting Systems," *British Medical Journal*, Vol. 320, pp. 759–763.

Bierly, P. E., and Spender, J. C. (1995), "Culture and High Reliability Organizations: The Case of the Nuclear Submarine," *Journal of Management*, Vol. 21, pp. 639–656.

Cao, C. G. L., and Milgram, P. (2000), "Disorientation in Minimal Access Surgery: A Case Study," in *Proceedings of the IEA 2000/HFES 2000 Congress*, San Diego, CA, Vol. 4, pp. 169–172.

Cao, C. G. L., and Taylor, H. (2004), "Effects of New Technology on the Operating Room Team," in *Work with Computing Systems, 2004*, H. M. Khalid, M. G. Helander, and A. W. Yeo, Eds., Damai Sciences, Kuala Lumpur, Malaysia, pp. 309–312.

Casey, S. (1993), *Set Phasers on Stun and Other True Tales of Design, Technology, and Human Error*, Aegean Park Press, Santa Barbara, CA.

CCPS (1992), *Guidelines for Hazard Evaluation Procedures, with Worked Examples*, 2nd ed., Center for Chemical Process Safety, American Institute of Chemical Engineers, New York.

CCPS (1994), *Guidelines for Preventing Human Error in Process Safety*, Center for Chemical Process Safety, American Institute of Chemical Engineers, New York.

Christoffersen, K., and Woods, D. D. (1999), "How Complex Human–Machine Systems Fail: Putting 'Human Error' in Context," in *The Occupational Ergonomics Handbook*, W. Karwowski and W. S. Marras, Eds., CRC Press, Boca Raton, FL, pp. 585–600.

Clark, H. H., and Schaefer, E. F. (1989), "Contributing to Discourse," *Cognitive Science*, Vol. 13, pp. 259–294.

Columbia Accident Investigation Board (2003), *Report Volume 1*, U.S. Government Printing Office, Washington, DC.

Cook, R. I., and Woods, D. D. (1996), "Adapting to New Technology in the Operating Room," *Human Factors*, Vol. 38, pp. 593–611.

Cook, R. I., Render, M., and Woods, D. D. (2000), "Gaps in the Continuity of Care and Progress on Patient Safety," *British Medical Journal*, Vol. 320, pp. 791–794.

Cullen, D. J., Bates, D. W., Small, S. D., Cooper, J. B., Nemeskal, A. R., and Leape, L. L. (1995), "The Incident Reporting System Does Not Detect Adverse Drug Events: A Problem for Quality Improvement," *Joint Commission Journal on Quality Improvement*, Vol. 21, pp. 541–548.

Damos, D. (1995), "Issues in Pilot Selection," in *Proceedings of the 8th International Symposium on Aviation Psychology*, Ohio State University, Columbus, OH, pp. 1365–1368.

Dekker, S. W. A. (2001), "The Disembodiment of Data in the Analysis of Human Factors Accidents," *Human Factors and Aerospace Safety*, Vol. 1, pp. 39–58.

Dekker, S. W. A. (2005), *Ten Questions About Human Error: A New View of Human Factors and System Safety*, Lawrence Erlbaum Associates, Mahwah, NJ.

Dey, A. K. (2001), "Understanding and Using Context," *Personal and Ubiquitous Computing*, Vol. 5, pp. 4–7.

Dhillon, B. S., and Singh, C. (1981), *Engineering Reliability: New Technologies and Applications*, Wiley, New York.

Drury, C. G. (1998), "Human Factors in Aviation Maintenance," in *Handbook of Aviation Human Factors*, D. J. Garland, J. A. Wise, and V. D. Hopkin, Eds., Lawrence Erlbaum Associates, Mahwah, NJ, pp. 591–606.

Embrey, D. E., Humphreys, P., Rosa, E. A., Kirwan, B., and Rea, K. (1984), *SLIM–MAUD: An Approach to Assessing Human Error Probabilities Using Structured Expert Judgment*, NUREG/CR-3518, U.S. Nuclear Regulatory Commission, Washington, DC.

Endsley, M. R. (1995), "Toward a Theory of Situation Awareness," *Human Factors*, Vol. 37, pp. 32–64.

Endsley, M. R., Bolté, B., and Jones, D. G. (2003), *Designing for Situation Awareness: An Approach to User-Centred Design*, CRC Press, Boca Raton, FL.

Evan, W. M., and Manion, M. (2002), *Minding the Machines: Preventing Technological Disasters*, Prentice-Hall, Upper Saddle River, NJ.

Fischhoff, B. (1975), "Hindsight–Foresight: The Effect of Outcome Knowledge on Judgment Under Uncertainty," *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 1, pp. 278–299.

Fraser, J. M., Smith, P. J., and Smith, J. W. (1992), "A Catalog of Errors," *International Journal of Man–Machine Systems*, Vol. 37, pp. 265–307.

Gertman, D. I., and Blackman, H. S. (1994), *Human Reliability and Safety Analysis Data Handbook*, Wiley, New York.

Grosjean, V., and Terrier, P. (1999), "Temporal Awareness: Pivotal in Performance?" *Ergonomics*, Vol. 42, pp. 443–456.

Hahn, A. H., and deVries J. A., II (1991), "Identification of Human Errors of Commission Using Sneak Analysis," in *Proceedings of the Human Factors Society 35th Annual Meeting*, pp. 1080–1084.

Helander, M. G. (1997), "The Human Factors Profession," in *Handbook of Human Factors and Ergonomics*, 2nd ed., G. Salvendy, Ed., Wiley, New York, pp. 3–16.

Hendrick, K., and Benner, L., Jr. (1987), *Investigating Accidents with STEP*, Marcel Dekker, New York.

Hofstede, G. (1991), *Cultures and Organizations: Software of the Mind*, McGraw-Hill, New York.

Hollnagel, E. (1993), *Human Reliability Analysis: Context and Control*, Academic Press, London.

Hollnagel, E. (1998), *Cognitive Reliability and Error Analysis Method*, Elsevier Science, New York.

Hollnagel, E., Ed. (2003), *Handbook of Cognitive Task Design*, Lawrence Erlbaum Associates, Mahwah, NJ.

Holroyd, C. B., and Coles, M. G. H. (2002), "The Neural Basis of Human Error Processing: Reinforcement Learning, Dopamine, and the Error-Related Negativity," *Psychological Review*, Vol. 109, pp. 679–709.

Johnson, W. G. (1980), *MORT Safety Assurance Systems*, Marcel Dekker, New York.

Johnson, C. (2002), "Software Tools to Support Incident Reporting in Safety-Critical Systems," *Safety Science*, Vol. 40, pp. 765–780.

Johnson, C. (2003), *Failure in Safety-Critical Systems: A Handbook of Accident and Incident Reporting*, University of Glasgow Press, Glasgow.

Kaber, D. B., and Endsley, M. R. (2004), "The Effects of Level of Automation and Adaptive Automation on Human Performance, Situation Awareness and Workload in a Dynamic Control Task," *Theoretical Issues in Ergonomics Science*, Vol. 4, pp. 113–153.

Kapur, K. C., and Lamberson, L. R. (1977), *Reliability in Engineering and Design*, Wiley, New York.

Kaye, K. (1999a), "Automated Flying Harbors Hidden Perils," *South Florida Sun-Sentinel*, September 27.

Kaye, K. (1999b), "United Has Eye in the Sky: Optical Recorders Check Crews," *South Florida Sun-Sentinel*, September 27.

Kaye, K. (1999c), "Program Urges Error Reporting, Mitigates Penalty," *South Florida Sun-Sentinel*, September 27.

Kepner, C. H., and Tregoe, B. B. (1981), *The New Rational Manager*, Kepner-Tregoe Inc., Princeton, NJ.

Kirwan, B. (1994), *A Guide to Practical Human Reliability Assessment*, Taylor & Francis, London.

Kirwan, B. (1999), "Some Developments in Human Reliability Assessment," in *The Occupational Ergonomics Handbook*, W. Karwowski and W. S. Marras, Eds., CRC Press, Boca Raton, FL, pp. 643–666.

Kirwan, B., and Ainsworth, L. K. (1992), *Guide to Task Analysis*, Taylor & Francis, London.

Kirwan, B., Martin, B. R., Rycraft, H., and Smith, A. (1990), "Human Error Data Collection and Data Generation," *International Journal of Quality and Reliability Management*, Vol. 7.4, pp. 34–66.

Kjellén, U. (2000), *Prevention of Accidents Through Experience Feedback*, Taylor & Francis, London.

Kohn, L. T., Corrigan, J. M., and Donaldson, M. S., Eds. (1999), *To Err Is Human: Building a Safer Health System*, National Academy Press, Washington, DC.

Koppel, R., Metlay, J. P., Cohen, A., Abaluck, B., Localio, A. R., Kimmel, S., and Strom, B. L. (2005), "Role of Computerized Physician Order Entry Systems in Facilitating Medication Errors," *Journal of the American Medical Association*, Vol. 293, pp. 1197–1203.

Kumamoto, H., and Henley, E. J. (1996), *Probabilistic Risk Assessment and Management for Engineers and Scientists*, 2nd ed., IEEE Press, Piscataway, NJ.

Leape, L. L., Brennan, T. A., Laird, N. M., Lawthers, A. G., Localio, A. R., Barnes, B. A., Hebert, L., Newhouse, J. P., Weiler, P. C., and Hiatt, H. H. (1991), "The Nature of Adverse Events in Hospitalized Patients: Results from the Harvard Medical Practice Study II," *New England Journal of Medicine*, Vol. 324, pp. 377–384.

Lee, J. D., and Moray, N. (1994), "Trust, Self-Confidence, and Operators' Adaptation to Automation," *International Journal of Human–Computer Studies*, Vol. 40, pp. 153–184.

Lee, J. D., and See, K. A. (2004), "Trust in Automation: Designing for Appropriate Reliance," *Human Factors*, Vol. 46, pp. 50–80.

Leech, D. S. (2004), "Learning in a Lean System," Defense Acquisition University Publication, Department of Defense, retrieved May 12, 2004, from http://acc.dau.mil/.

Lekberg, A. (1997), "Different Approaches to Accident Investigation: How the Analyst Makes the Difference," in *Proceedings of the 15th International Systems Safety Conference*, Sterling, VA, International Systems Safety Society, pp. 178–193.

Levy, J., Gopher, D., and Donchin, Y. (2002), "An Analysis of Work Activity in the Operating Room: Applying Psychological Theory to Lower the Likelihood of Human Error," in *Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting*, Human Factors and Ergonomics Society, Santa Monica, CA, pp. 1457–1461.

Luczak, H. (1997), "Task Analysis," in *Handbook of Human Factors and Ergonomics*, 2nd ed., G. Salvendy, Ed., Wiley, New York, pp. 340–416.

Moray, N. (2000), "Culture, Politics and Ergonomics," *Ergonomics*, Vol. 43, pp. 858–868.

Moray, N., Inagaki, T., and Itoh, M. (2000), "Adaptive Automation, Trust, and Self-Confidence in Fault Management of Time-Critical Tasks," *Journal of Experimental Psychology: Applied*, Vol. 6, pp. 44–58.

Nielsen, J. (1995), *Usability Engineering*, Academic Press, San Diego, CA.

Nieuwenhuis, S. N., Ridderinkhof, K. R., Blom, J., Band, G. P. H., and Kok, A. (2001), "Error Related Brain Potentials Are Differentially Related to Awareness of Response Errors: Evidence from an Antisaccade Task," *Psychophysiology*, Vol. 38, pp. 752–760.

Norman, D. A. (1981), "Categorization of Action Slips," *Psychological Review*, Vol. 88, pp. 1–15.

Parasuraman, R. (2003), "Neuroergonomics: Research and Practice," *Theoretical Issues in Ergonomics Science*, Vol. 4, pp. 5–20.

Parasuraman, R., and Riley, V. (1997), "Humans and Automation: Use, Misuse, Disuse, and Abuse," *Human Factors*, Vol. 39, pp. 230–253.

Parasuraman, R., Sheridan, T. B., and Wickens, C. D. (2000), "A Model for Types and Levels of Human Interaction with Automation," *IEEE Transactions on Systems Man, and Cybernetics, Part A: Systems and Humans*, Vol. 30, pp. 276–297.

Perrow, C. (1983), "The Organizational Context of Human Factors Engineering," *Administrative Science Quarterly*, Vol. 27, pp. 521–541.

Perrow, C. (1999), *Normal Accidents: Living with High-Risk Technologies*, Princeton University Press, Princeton, NJ.

Phillips, L. D., Embrey, D. E., Humphreys, P., and Selby, D. L. (1990), "A Sociotechnical Approach to Assessing Human Reliability," in *Influence Diagrams, Belief Nets and Decision Making: Their Influence on Safety and Reliability*, R. M. Oliver and J. A. Smith, Eds., Wiley, New York.

Potter, S. S., Roth, E. M., Woods, D. D., and Elm, W. C. (1998), "A Framework for Integrating Cognitive Task Analysis into the System Development Process," in *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting*, Human Factors and Ergonomics Society, Santa Monica, CA, pp. 395–399.

Prager, L. O. (1998), "Sign Here," *American Medical News*, Vol. 41, October 12, pp. 13–14.

Rasmussen, J. (1982), "Human Errors: A Taxonomy for Describing Human Malfunction in Industrial Installations," *Journal of Occupational Accidents*, Vol. 4, pp. 311–333.

Rasmussen, J. (1986), *Information Processing and Human–Machine Interaction: An Approach to Cognitive Engineering*, Elsevier, New York.

Rasmussen, J., Pejterson, A. M., and Goodstein, L. P. (1994), *Cognitive Systems Engineering*, Wiley, New York.

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

Reason, J. (1997), *Managing the Risks of Organizational Accidents*, Ashgate, Aldershot, Hampshire, England.

Reason, J., and Hobbs, A. (2003), *Managing Maintenance Error: A Practical Guide*, Ashgate, Aldershot, Hampshire, England.

Reynolds-Mozrall, J., Drury, C. G., Sharit, J., and Cerny, F. (2000), "The Effects of Whole-Body Restriction on Task Performance," *Ergonomics*, Vol. 43, pp. 1805–1823.

Roberts, K. H. (1990), "Some Characteristics of One Type of High Reliability Organization," *Organization Science*, Vol. 1, pp. 160–176.

Robinson, A. G., and Stern, S. (1998), *Corporate Creativity*, Berrett-Koehler, San Francisco.

Rochlin, G., La Porte, T. D., and Roberts, K. H. (1987), "The Self-Designing High Reliability Organization: Aircraft Carrier Flight Operations at Sea," *Naval War College Review*, Vol. 40, pp. 76–90.

Rosenthal, J., Booth, M., and Barry, A. (2001), "Cost Implications of State Medical Error Reporting Programs: A Briefing Paper," National Academy for State Health Policy, Portland, ME.

Rouse, W. B., and Rouse, S. (1983), "Analysis and Classification of Human Error," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-13, pp. 539–549.

Rumelhart, D. E., and McClelland, J. L., Eds. (1986), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1, *Foundations*, MIT Press, Cambridge, MA.

Sanders, M. S., and McCormick, E. J. (1993), *Human Factors in Engineering and Design*, 7th ed., McGraw-Hill, New York.

Senders, J. W., and Moray, N. P. (1991), *Human Error: Cause, Prediction, and Reduction*, Lawrence Erlbaum Associates, Mahwah, NJ.

Sharit, J. (1997), "Allocation of Functions," in *Handbook of Human Factors and Ergonomics*, 2nd ed., G. Salvendy, Ed., Wiley, New York, pp. 301–339.

Sharit, J. (1998), "Applying Human and System Reliability Analysis to the Design and Analysis of Written Procedures in High-Risk Industries," *Human Factors and Ergonomics in Manufacturing*, Vol. 8, pp. 265–281.

Sharit, J. (2003), "Perspectives on Computer Aiding in Cognitive Work Domains: Toward Predictions of Effectiveness and Use," *Ergonomics*, Vol. 46, pp. 126–140.

Shepherd, A. (2000), *Hierarchical Task Analysis*, Taylor & Francis, London.

Simon, H. A. (1966), *Models of Man: Social and Rational*, Wiley, New York.

Smith, P. J., and Geddes, N. D. (2003), "A Cognitive Systems Engineering Approach to the Design of Decision Support Systems," in *The Human–Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*, J. A. Jacko and A. Sears, Eds., Lawrence Erlbaum Associates, Mahwah, NJ.

Stout, R. M., Cannon-Bowers, J. A., Salas, E., and Milanovich, D. M. (1999), "Planning, Shared Mental Models, and Coordinated Performance: An Empirical Link Is Established," *Human Factors*, Vol. 41, pp. 61–71.

Strauch, B. (2002), *Investigating Human Error: Incidents, Accidents, and Complex Systems*, Ashgate, Aldershot, Hampshire, England.

Swain, A. D., and Guttmann, H. E. (1983), *Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications*, NUREG/CR-1278, U.S. Nuclear Regulatory Commission, Washington, DC.

Thomas, E. J., and Helmreich, R. L. (2002), "Will Airline Safety Models Work in Medicine?" in *Medical Error: What Do We Know? What Do We Do?* M. M. Rosenthal and K. M. Sutcliffe, Eds., Jossey-Bass, San Francisco, pp. 217–234.

Turrell, M. (2002), "Idea Management and the Suggestion Box," White Paper, Imaginatik Research, retrieved May

12, 2004, from http://www.imaginatik.com/web/nsf/docs/idea_reports_imaginatik.

Vicente, K. J. (1999), *Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work*, Lawrence Erlbaum Associates, Mahwah, NJ.

Vicente, K. J. (2004), *The Human Factor: Revolutionizing the Way People Live with Technology*, Routledge, New York.

Weiner, E. L. (1985), "Beyond the Sterile Cockpit," *Human Factors*, Vol. 27, pp. 75–90.

Wickens, C. D. (1984), "Processing Resources in Attention," in *Varieties of Attention*, R. Parasuraman and R. Davies, Eds., Academic Press, New York, pp. 63–101.

Wickens, C. D., Liu, Y., Becker, S. E. G., and Lee, J. D. (2004), *An Introduction to Human Factors Engineering*, 2nd ed., Prentice-Hall, Upper Saddle River, NJ.

Wilde, G. J. S. (1982), "The Theory of Risk Homeostasis: Implications for Safety and Health," *Risk Analysis*, Vol. 2, pp. 209–225.

Woods, D. D. (1984), "Some Results on Operator Performance in Emergency Events," *Institute of Chemical Engineers Symposium Series*, Vol. 90, pp. 21–31.

Woods, D. D. (1993), "Process Tracing Methods for the Study of Cognition Outside the Experimental Psychology Laboratory," in *Decision Making in Action: Models and Methods*, G. Klein, R. Calderwood, and J. Orasanu, Eds., Ablex, Norwood, NJ, pp. 227–251.

Woods, D. D., and Watts, J. C. (1997), "How Not to Navigate Through Too Many Displays," in *Handbook of Human–Computer Interaction*, 2nd ed., M. Helander, T. K. Landauer, and P. Prabhu, Eds., Elsevier Science, New York, pp. 617–650.

Woods, D. D., Johannesen, L. J., Cook, R. I., and Sarter, N. B. (1994), "Behind Human Error: Cognitive Systems, Computers, and Hindsight," CSERIAC State-of-the Art-Report, Crew Systems Ergonomics Information Analysis Center, Wright-Patterson Air Force Base, OH.

Woods, D. D., Sarter, N. B., and Billings, C. E. (1997), "Automation Surprises," in *Handbook of Human Factors and Ergonomics*, 2nd ed., G. Salvendy, Ed., Wiley, New York, pp. 1926–1943.