SPECIAL SECTION

Doctors Are More Dangerous Than Gun Owners: A Rejoinder to Error Counting

Sidney W. A. Dekker, Lund University, Ljungbyhed, Sweden

Objective: This paper analyzes some of the problems with error counting as well as the difficulty of proposing viable alternatives. **Background:** Counting and tabulating negatives (e.g., errors) are currently popular ways to measure and help improve safety in a variety of domains. They uphold an illusion of rationality and control but may offer neither real insight nor productive routes for improving safety. **Method:** The paper conducts a critical analysis of assumptions underlying error counting in human factors. **Results:** Error counting is a form of structural analysis that focuses on (supposed) causes and consequences; it defines risk and safety instrumentally in terms of minimizing negatives and their measurable effects. In this way, physicians can be proven to be 7500 times less safe than gun owners, as they are responsible for many more accidental deaths. **Conclusion:** The appeal of error counting may lie in a naive realism that can enchant researchers and practitioners alike. Supporting facts will continue to be found by those looking for errors through increasingly refined methods. **Application:** The paper outlines a different approach to understanding safety in complex systems that is more socially and politically oriented and that places emphasis on interpretation and social construction rather than on putatively objective structural features.

INTRODUCTION

There are about 700,000 physicians in the United States. The U.S. Institute of Medicine estimates that each year between 44,000 and 98,000 people die as a result of medical errors (Kohn, Corrigan, & Donaldson, 1999). This makes for a yearly accidental death rate per doctor of between 0.063 and 0.14. In other words, up to one in seven doctors will kill a patient each year by mistake. In contrast, there are 80 million gun owners in the United States. They are responsible for 1,500 accidental gun deaths in a typical year (e.g., National Safety Council, 2004). This means that the accidental death rate, caused by gun owner error, is 0.000019 per gun owner per year. Only about 1 in 53,000 gun owners will kill somebody by mistake. Doctors, then, are 7,500 times more likely than gun owners to kill somebody as a result of human error (Dekker, 2005).

Although the comparison between doctors and gun owners is ridiculous for many reasons, and is meant facetiously, organizations and other stakeholders (e.g., consumers, trade and industry groups, regulators, researchers) actually do use error counts in trying to assess the “safety health” of an organization or professional group. This would seem to carry many advantages. Not only does error counting provide an immediate, numeric estimate of the probability of accidental death, injury, or other undesirable event, it also allows comparison (this hospital vs. that hospital, this airline vs. that one). Keeping track of adverse events is thought to provide relatively easy, quick, and accurate access to the internal safety workings of a system. Adverse events can be seen as the start of, or reason for, deeper probing to search for environmental threats or unfavorable conditions that could be changed to prevent recurrence.

Over the past three decades, human factors researchers have spawned a number of error classification systems. Some classify decision errors together with the conditions that helped produce them (Kowalsky, Masters, Stone, Babcock, & Rypka,

Address correspondence to Sidney W. A. Dekker, Lund University, School of Aviation, 260 70 Ljungbyhed, Sweden; sidney.dekker@tfhs.lu.se. ***HUMAN FACTORS,*** Vol. 49, No. 2, April 2007, pp. 177–184.

1974). Some have a specific goal. For example, they aim to categorize information transfer problems that may happen during instructions, watch changeover briefings, or other coordination (Billings & Cheaney, 1981). Others try to divide error causes into cognitive, social, and situational (physical/environmental/ergonomic) factors (Fegetter, 1982). Yet others attempt to classify error causes along the lines of a linear information-processing/decision-making model (Rouse & Rouse, 1983). Various counting methods are founded on their own models (e.g., threat and error model; Helmreich, Klinect, & Wilhelm, 1999), whereas others apply, for example, the Swiss-cheese metaphor in the search for errors and vulnerabilities up the causal chain (Shappell & Wiegmann, 2001). This metaphor suggests that systems have multiple layers of defense but all of them have holes, which need to line up to allow an accident (see Reason, 1990).

In the categorization and tabulation of errors, researchers make a number of assumptions and take certain philosophical positions. Few of these are made explicit in the description of these methods, yet they carry consequences for the utility and quality of the error count as a measure of safety health and as a tool for directing resources for improvement. In this paper, I will examine some of those assumptions, including the naively realist idea that social phenomena (including errors) exist as facts outside of individual minds, open for objective scrutiny by anybody with an appropriate method. I will show some fairly obvious counterinstances of this assumption but then acknowledge that moving human factors away from this idea is extremely difficult because observed facts always privilege the ruling paradigm. I nonetheless conclude by making a proposal for a new standard in which the assumption is no longer that safety, once established, can be maintained by requiring human performance to stay within the prespecified boundaries of an error categorization tool. Instead, I argue for the development of better ways to understand how people and organizations themselves create safety through practice. I also argue for greater self-consciousness on the part of researchers and other stakeholders: How well calibrated are the models of safety and risk that are expressed through existing methods and proposed countermeasures? After all, the models are but instances, all negotiable and refutable, of an inherently and permanently imperfect knowledge base of what makes systems brittle or resilient.

Errors Exist "Out There" and Can Be Discovered With a Good Method

Error counting generally assumes that there is a reality "out there" that researchers should try to approach as closely as possible. For this, they need a good method. This is a firmly modernist stance, one that has dominated science for centuries. Errors, in this sense, are a kind of Durkheimian fact (Durkheim, 1895/1950). Reality exists; the truth can be found. A scientifically based method helps people do just that, as it supposedly eliminates subjective preconceptions and enables people to know reality just as it is.

But when it comes to errors, this turns out to be complicated. What, for example, causes errors? Having an idea about their cause is often crucial for the ability to categorize errors using one of the methods mentioned previously. According to Helmreich (2000), "errors result from physiological and psychological limitations of humans. Causes of error include fatigue, workload, and fear, as well as cognitive overload, poor interpersonal communications, imperfect information processing, and flawed decision making" (p. 746). But this is circular: Do errors cause flawed decision making, or does flawed decision making lead to errors? The objective observation of errors is suddenly no longer so simple. Mixing up cause and consequence is typical for error categorization methods (Dougherty, 1990; Hollnagel, 1998), but to their adherents, such causal confounds are neither really surprising nor really problematic. Truth, after all, can be elusive. What matters is getting the method right. More method will presumably solve problems of method.

Other problems supposedly related to method also occur. In a classification scheme currently popular in aviation, Line-Oriented Safety Audits (see Helmreich et al., 1999), the observer is asked to distinguish, among other things, between "procedure errors" and "proficiency errors." Proficiency errors are related to a lack of skills, experience, or (recent) practice, whereas procedure errors are those that occur while carrying out prescribed or normative sequences of action (e.g., checklists). This seems straightforward. But, as Croft (2001) reported, the following problem confronts the observer. One type of error (a pilot entering a wrong altitude in a flight computer) can legitimately end up in either of the two categories: "For example, entering the wrong flight altitude in the flight management system is considered a procedural error. . . . Not knowing how to use certain automated features in an aircraft's flight computer is considered a proficiency error" (Croft, 2001, p. 77).

If a pilot enters the wrong flight altitude in the flight management system, is that a procedural or a proficiency issue? If there are problems in matching observed facts with theory (e.g., one factual observation can comfortably fit two categories), then researchers typically see these as problems of method, calling for further refinement. For example, the measuring instruments can be made more sensitive, so that they discriminate better between different observations. Observers can also be trained better, so that they recognize subtle differences between errors and learn to code them correctly. These are typical responses of a research community to the challenges raised by mismatches between theory and observed fact.

But are these problems of method? This is the crucial question. Kuhn (1962) encouraged science to turn to creative philosophy when confronted with the inklings of problems in relating theory to observations. It can be an effective way to elucidate and, if necessary, weaken the grip of a tradition upon the collective mind. It may even suggest the basis for a new direction. For any scientific endeavor, such reconsideration is appropriate when epistemological questions arise – questions about how people know what they (think they) know.

Is It an Error? That Depends on Who You Ask

Consider a study reported by Hollnagel and Amalberti (2001), whose purpose was to test an error measurement instrument. The method asked observers to count errors and categorize errors using a taxonomy proposed by the developers. It was tested in a field setting by pairs of psychologists and air traffic controllers who studied air traffic control work in real time. Despite common indoctrination, there were substantial differences between the numbers and kinds of errors each of the two groups of observers noted, and only a very small number of errors were observed by both. Air traffic controllers relied on external working conditions (e.g., interfaces, personnel, and time resources) to refer to and categorize errors, whereas psychologists preferred to locate the error somewhere in presumed quarters of the mind (e.g., working memory) or in some mental state (e.g., attentional lapses).

Moreover, air traffic controllers who actually did the work could tell the error coders that they both had it wrong. Observed "errors" were not errors to those "committing" them but, rather, deliberate strategies intended to manage problems or foreseen situations that the error counters had neither seen nor understood as such if they had. Such normalization of actions, which at first appear deviant from the outside, is a critical aspect of understanding human work and its strengths and weaknesses (see Vaughan, 1999). Croft (2001) reported the same result in cockpits: More than half the "errors" revealed by error counters were never discovered by the flight crews themselves. Some realists may argue that the ability to discover errors not seen by people themselves confirms the superiority of the method. But such claims of epistemological privilege are hubris. As Jones (1986) pointed out, trying to study a social phenomenon, such as error, independent of meanings attached to it runs the risk of abstracting some essentialist definition of error that bears no relation to the practices and interpretations in question. In addition, it runs the risk of unconsciously imposing one's own subjective interpretation under the guise of detached, scientific observation.

Error Counting and Naive Realism

At first sight, Hollnagel's and Amalberti's air traffic control study raises the question of whose standard is right. If there is disagreement about what an observation means (i.e., whether it is an error or not), the question becomes one of arbitrage. Who can make the strongest epistemological claim? Many people would probably put their bet on the practitioner. But this misses the point. If particular observers describe reality in a particular way (e.g., this was a "procedure error"), then that does not imply any type of mapping onto an objectively attainable external reality – close or remote, good or bad. Postmodernists argue that a single, stable reality that can be most closely approximated by the best method or the most qualified observer does not exist (Capra, 1982). Although people seem to need the idea of a fixed, stable reality surrounding them, independent of who looks at it, the foregoing example denies them this.

The reality of an observation is socially constructed. The error becomes true (or appears to people as a close correspondence to some objective reality) only because a community of specialists has developed tools that would seem to make it appear and have agreed on the language that makes it visible. There is nothing inherently "true" about the error at all. Its meaning is merely enforced and handed down through systems of observer training, labeling and communication of the results, and industry acceptance and promotion.

Observed Facts Are Created by the Method Itself

Even though an observed error may appear as entirely real and "factual" to the observer, that does not mean that it is. Facts privilege the ruling paradigm. Facts actually exist by virtue of the current paradigm. They can be neither discovered nor given meaning without it. The autonomy principle is false: Facts that are available as objective content of one theory are not equally available to another, as the theory itself helps construct them: "On closer analysis, we even find that science knows no 'bare facts' at all, but that the 'facts' that enter our knowledge are already viewed in a certain way and are, therefore, essentially ideational" (Feyerabend, 1993, p. 11).

Researchers who apply a theory of naturalistic decision making, for example, will not see a "procedure error." They may instead see a continuous control task, a flow of actions and assessments, coupled and mutually cued – a flow with nonlinear feedback loops and interactions, inextricably embedded in a multilayered evolving context. Such a characterization is hostile to the digitization necessary to fish out individual "human errors." Observers are themselves participants, participating in the very creation of the observed fact. (Even in a crude sense this would be true: Observing performance probably distorts people's normal practice, perhaps turning situated performance into window-dressed posture.)

STANDING FIRM: THE THEORY IS RIGHT

Kuhn (1962) resisted the idea that science progresses through the accumulation of observed facts that disagree with, and ultimately manage to topple, a theory. Counterinstances are seen only as further puzzles in the match between observation and theory, to be addressed by more method. It is extremely difficult for communities to renounce the paradigm that has led them into a crisis. Instead, epistemological difficulties suffered by error-counting methods (was this a cause or consequence, a procedural or proficiency error?) are entertained as reasons to engage in yet more methodological refinement consonant with the current paradigm. It can adopt a kind of self-sustaining energy, or "consensus authority" (see Angell & Straub, 1999), in which nobody questions error counting because everybody is doing it. In accepting the utility of error counting, it is likely that industry accepts its theory (and thereby the reality and validity of the observations it generates) on the authority of authors, teachers, and their texts, not because of evidence. Croft's 2001 headline in Aviation Week & Space Technology announced, "Researchers perfect new ways to monitor pilot performance." If researchers have perfected a method, there is little an industry can do other than accept such authority. What alternatives have they, Kuhn (1962) would ask, or what competence?

Nobody is willing to forgo a paradigm until and unless a viable alternative is ready to take its place. This is a sustained argument for the continuation of error counting: Researchers are willing to acknowledge that what they do is not perfect but vow to keep going until shown something better, and industry concurs. As Kuhn (1962) would say, the decision to reject one paradigm necessarily coincides with the embrace of another.

The Difficulty of Proposing an Alternative Theory

Proposing a viable alternative theory that can assimilate its own facts, however, is exceedingly difficult. Facts, after all, privilege the status quo. Galileo's telescopic observations of the sky motivated an alternative explanation about the place of the earth in the universe, which favored the Copernican heliocentric interpretation (in which Earth goes around the Sun) over the Ptolomeic geocentric one. The Copernican interpretation, however, was a worldview away from the ruling interpretation, and many doubted Galileo's data as a valid empirical window on a heliocentric universe. People were suspicious of the new instrument. Some asked Galileo to open up his telescope to prove that there was no little moon hiding inside of it (Feyerabend, 1993).

One problem was that Galileo did not offer a theory for why the telescope was supposed to offer a better picture of the sky than the naked eye. He could not, because relevant concepts (optica) were not yet well developed. Generating better data (as Galileo did) and developing new methods for better access to these data (e.g., a telescope) does in itself little to dislodge an established theory that allows people to see a phenomenon with their naked eye and explain it with their common sense. The Sun goes around Earth. Earth is fixed. The Church was right, and Galileo was wrong. None of the observed facts could prove him right because there was no coherent set of theories ready to accommodate his facts and give them meaning. The Church was right, as *it* had all the facts – and it had the theory to assimilate them.

Interestingly, the Church kept closer to reason as it was defined at the time. It considered the social, political, and ethical implications of Galileo's alternatives and deemed them too risky to accept. Disavowing the geocentric idea would be disavowing creation itself, removing *the* common ontological denominator of the past millennium and severely undermining the authority and political power the Church derived from it. Error classification methods, too, guard a rationality that many would hate to see disintegrate. Without errors, without such a "factual" basis, how could one hold people accountable for mistakes or report safety occurrences and maintain expensive incident reporting schemes? What could people fix if there are no "causes"? They should, rather, hold onto the realist status quo and cause minimal disruption to the existing theory. And they can, for most observed facts still seem to privilege it. Errors exist. They must.

If You Cannot See Errors, You Are Not a Good Psychologist

To the naive realist, the argument that errors exist is not only natural and necessary – it is also quite impeccable. The idea that errors do not exist, in contrast, is unnatural. It is absurd. Those within the established paradigm will challenge the legitimacy of questions raised about the existence of errors and the legitimacy of those who raise the questions: "Indeed, there are some psychologists who would deny the existence of errors altogether. We will not pursue that doubtful line of argument here" (Reason & Hobbs, 2003, p. 39).

If some scientists do not succeed in bringing statement and fact into closer agreement (they do not see a "procedure error" where others would), then this discredits the scientist rather than the theory. Galileo suffered from this, too. It was the scientist who was discredited (for a while, at least), not the prevailing paradigm. So what did he do? Galileo engaged in propaganda and psychological trickery (Feyerabend, 1993). Through imaginary conversations among Sagredo, Salviati, and Simplicio, written in his native Italian rather than Latin, he put the ontological uncertainty and epistemological difficulty of the geocentric interpretation on full display. Where the appeal to empirical facts fails, an appeal to logic may still succeed. The same is true for error counting and classification. Just imagine this dialogue (see Dekker, 2005, p. 58–59):

> Simplicio: "Errors result from physiological and psychological limitations of humans. Causes of error include fatigue, workload, and fear, as well as cognitive overload, poor interpersonal communications, imperfect information processing, and flawed decision making."
> Sagredo: "But are errors in this case not simply the result of other errors? Flawed decision making would be an error. But in your logic, it causes an error. What is the 'error' then? And how can we categorize it?"
> Simplicio: "Well, but errors are caused by poor decisions, failures to adhere to instructions, failures to prioritize attention, improper procedure, and so forth."
> Sagredo: "This appears to be not causal explanation, but simply relabeling. Whether you say 'error,' or 'poor decision,' or 'failure to prioritize attention,' it all still sounds like 'error,' at least when interpreted in your worldview. And how can one be the cause of the other to the exclusion of the other way around? Can 'errors' cause 'poor decisions' just like 'poor decisions' cause 'errors'? There is nothing in your logic that rules this out, but then we end up with a tautology, not an explanation."

And yet, such arguments may not help, either. The appeal to logic may fail in the face of overwhelming support for a ruling paradigm – support that derives from consensus authority. Even Einstein expressed amazement at the common reflex to rely on measurements (e.g., error counts) rather than logic and argument: "Is it not really strange," Albert Einstein asked in a letter to Max Born (quoted in Feyerabend, 1993, p. 239), "that human beings are normally deaf to the strongest of argument while they are always inclined to overestimate measuring accuracies?"

Numbers are strong. Arguments are weak. Error counting is good because it generates numbers. It relies on putatively accurate measurements (recall Croft, 2001: "Researchers" have "perfected" ways to monitor pilot performance). People will reject no theory on the basis of argument or logic alone. They need another to take its place.

ABANDON THE IDEA OF ERRATIC PEOPLE IN SAFE SYSTEMS

The dominant safety paradigm in human factors has long been based on searching for ways to limit human variability in otherwise safe systems (Hollnagel, Woods, & Leveson, 2006; Woods, Johannesen, Cook, & Sarter, 1994). The assumption is that safety, once established, can be maintained by requiring human performance to stay within prescribed boundaries. Error counting and categorizing operationalizes this assumption by trying to observe how performance deviates from, or strays outside, established norms (e.g., violations of procedure, inadequate proficiency). Indeed, error counting assumes that the quantity measured (errors) has a meaningful relationship with the quality investigated (safety). It goes without saying that more of the quantity gives less of the quality. Such a connection is a folk model, at best, and is actually unsupported by evidence. Instead, studies of how complex systems succeed, and sometimes fail, demonstrate a much more complex, and much less instrumental, relationship among external worlds of cause and effect, social worlds of human relationships, and inner worlds of values and meaning.

The formal descriptions of work embodied in policies, procedures, and regulations – and implicitly imposed through error counting – are incomplete as models of expertise and success (e.g., Hollnagel et al., 2006). In a world of finite resources, uncertainty, and multiple conflicting goals, the knowledge base for generating safety in complex, risky operations is inherently and permanently imperfect (Rochlin, 1999), and no externally dictated logics of an error categorization system can arbitrate in any lasting way between what is safe and what is unsafe. The issue is not, therefore, whether potentially erratic human performance stays within or strays outside the perimeters of artificially imposed error categories, for those categories themselves represent only a particular slice of the knowledge base, or a particular model of risk, about what makes operations resilient or brittle. This representation is probably an obsolete, coarse approximation at best. There are two interesting issues: The first is how practitioners themselves continually contribute to the creation of safety through their practice at all levels of an organization and how self-conscious these practitioners are with respect to those constructions of safety and risk. The second is how researchers and other stakeholders develop and sustain the models of risk that find their expression in the methods and countermeasures they deploy, and whether these stakeholders are sufficiently self-conscious to acknowledge that those models may be ill calibrated, or a bad fit, and ready for reconsideration and renewal. In other words, do researchers and stakeholders themselves monitor, and critically question, how they monitor safety? In conclusion, I turn to these two issues now.

People Create Safety Through Practice

Where the creation of safety appears to have everything to do with people learning about, and adapting around, multiple goals, hazards, and trade-offs, deeper investigation of most stories of "error" show that failures represent breakdowns in adaptations directed at coping with such complexity (e.g., Cook, 1998; Hollnagel et al., 2006; Rochlin, 1999). Among other things, they indicate the following:

- Practitioners and organizations continually assess and revise their approaches to work in an attempt to remain sensitive to the possibility of failure. Efforts to create safety, in other words, are ongoing. Not being successful is related to limits of the current model of competence and, in a learning organization, reflects a discovery of those boundaries.
- Strategies that practitioners and organizations maintain for coping with potential pathways to failure can either be strong or resilient (i.e., well calibrated) or weak and mistaken (i.e., ill calibrated).
- Organizations and people can also become overconfident in how well calibrated their strategies are. Effective organizations remain alert for signs that circumstances exist, or are developing, in which that confidence is erroneous or misplaced (Gras, Moricot, Poirot-Delpech, & Scardigli, 1990/1994; Rochlin, 1993). This, after all, can avoid narrow interpretations of risk and stale countermeasures.

Safe operation, accordingly, has little to do with the structural descriptors sought by error counts ("violations," "proficiency errors"), nor is safety the instrumental outcome of a minimization of errors and their presumably measurable effects. Safety does not exist "out there," independent of people's minds or culture, ready to be measured by looking at behavior alone (Slovic, 1992). Instead, insight has been growing that research into safe operations should consider safety as a dynamic, interactive, communicative act that is created as people conduct work, construct discourse and rationality around it, and gather experiences from it (e.g., Orasanu, 2001). Cultures of safety are not cultures without errors or violations – on the contrary. Practitioners are not merely in the business of managing risk or avoiding error, if they are that at all. Rather, they actively engage operational and organizational conditions to intersubjectively construct their beliefs in the possibility of continued operational safety. This includes anticipation of events that could have led to serious outcomes, complemented by the continuing expectation of future surprise. "Safety is in some sense a story a group or organization tells about itself and its relation to its task environment" (Rochlin, 1999, p. 1555).

Particular aspects of how organization members tell or evaluate safety stories can serve as markers (see Columbia Accident Investigation Board, 2003). In *Creating Foresight* (2003), Woods (p. 5), for example, called one of these "distancing through differencing." In this process, organizational members look at other failures and other organizations as not relevant to them and their situation. They discard other events because they appear to be dissimilar or distant. Discovering this through qualitative inquiry can help specify how people and organizations reflexively create their idea, their story of safety. Just because the organization or section has different technical problems, different managers, different histories, or can claim to already have addressed a particular safety concern revealed by the event does not mean that they are immune to the problem. Seemingly divergent events can represent similar underlying patterns in the drift toward hazard. High-reliability organizations characterize themselves through their preoccupation with failure: They continually ask themselves how things can go wrong or could have gone wrong, rather than congratulating themselves that things went right. Distancing through differencing means underplaying this preoccupation. It is one way to prevent learning from events elsewhere, one way to throw up obstacles in the flow of safety-related information.

Additional processes that can be discovered include the extent to which an organization resists oversimplifying interpretations of operational data – whether it defers to expertise and expert judgment rather than managerial imperatives, and whether it sees continued operational success as a guarantee of future safety, as an indication that hazards are not present or that countermeasures in place suffice. Also, it could be interesting to probe to what extent problem-solving processes are disjointed across organizational departments, sections, or subcontractors, as discontinuities and internal handovers of tasks increase risk (Vaughan, 1999). With information incomplete, disjointed, and patchy, nobody may be able to recognize the gradual erosion of safety constraints on the design and operation of the original system.

Monitoring How Safety Is Monitored

It is, of course, a matter of debate whether the higher order organizational processes that could be part of new safety probes (e.g., distancing through differencing, deference to expertise, fragmentation of problem solving, incremental judgments into disaster) are any more real than the errors from the counting methods they seek to replace or augment. But then, the reality of these phenomena is in the eye of the beholder. The processes and phenomena are real enough to those who look for them and who wield the theories to accommodate the results. Criteria for success may lie elsewhere – for example, in how well the measure maps onto past evidence of precursors to failure. Yet even such mappings are subject to paradigmatic interpretations of the evidence base. Indeed, consonant with the ontological relativity of the age human factors has now entered, the debate can probably never be closed. Are doctors more dangerous than gun owners? Do errors exist? It depends on whom you ask.

The real issue, therefore, lies a step away from the fray. A level up, if you will. Whether errors are counted as Durkheimian fact or safety is seen as a reflexive project, competing premises and practices reflect particular models of risk. These models of risk are interesting not because of their differential abilities to access empirical truth (because that may all be relative) but because of what they say about the creators or proponents of the models. It is not merely the monitoring of safety that should be pursued but the monitoring of that monitoring (*Creating Foresight,* 2003). To make progress in safety, one important step is to engage in such meta-monitoring. Researchers should become better aware of the models of risk embodied in their approaches to safety. Whether doctors are more dangerous than gun owners, in other words, is irrelevant. What matters is what the respective communities see as their dominant sources of risk and how that, in turn, informs the measures and countermeasures they apply.

REFERENCES

Angell, I. O., & Straub, B. (1999). Rain-dancing with pseudo-science. *Cognition, Technology and Work, 1,* 179–196.

Billings, C. E., & Cheaney, E. S. (1981). *Information transfer problems in the aviation system* (NASA Tech. Paper 1875). Moffett Field, CA: NASA Ames Research Center.

Capra, F. (1982). *The turning point.* New York: Simon and Schuster.

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

Cook, R. I. (1998). Two years before the mast: Learning how to learn about patient safety. In *Proceedings of Enhancing Patient Safety and Reducing Errors in Health Care: A Multidisciplinary Leadership Conference* (pp. 61–65). Rancho Mirage, CA: Annenberg Center for Health Sciences at Eisenhower.

Creating foresight: How resilience engineering can transform NASA's approach to risky decision making: U.S. Senate testimony for the Committee on Commerce, Science and Transportation, 108th Cong., Future of NASA SR-253 (2003) (testimony of David Woods, Professor, Institute for Ergonomics, The Ohio State University).

Croft, J. (2001). Researchers perfect new ways to monitor pilot performance. *Aviation Week and Space Technology, 155(3),* pp. 76–77.

Dekker, S. W. A. (2005). *Ten questions about human error: A new view of human factors and system safety.* Mahwah, NJ: Erlbaum.

Dougherty, E. M., Jr. (1990). Human reliability analysis: Where shouldst thou turn? *Reliability Engineering and System Safety, 29,* 283–299.

Durkheim, E. (1950). *The rules of the sociological method*. New York: Free Press. (Original work published 1895)

Fegetter, A. J. (1982). A method for investigating human factors aspects of aircraft accidents and incidents. *Ergonomics, 25,* 1065–1075.

Feyerabend, P. (1993). *Against method* (3rd ed.). London: Verso.

Gras, A., Moricot, C., Poirot-Delpech, S. L., & Scardigli, V. (1994). *Faced with automation: The pilot, the controller, and the engineer* (J. Lundsten, Trans.). Paris: Publications de la Sorbonne. (Original work published 1990)

Helmreich, R. L. (2000). On error management: Lessons from aviation. *British Medical Journal, 320,* 745–753.

Helmreich, R. L., Klinect, J. R., & Wilhelm, J. A. (1999). Models of threat, error and CRM in flight operations. In R. S. Jensen (Ed.), *Proceedings of the 10th International Symposium on Aviation Psychology* (pp. 677–682). Columbus: Ohio State University.

Hollnagel, E. (1998). *Cognitive reliability and error analysis method (CREAM).* Oxford, UK: Elsevier Science.

Hollnagel, E., & Amalberti, R. (2001). The emperor's new clothes: Or whatever happened to "human error"? In S. W. A. Dekker (Ed.), *Proceedings of the 4th International Workshop on Human Error, Safety and Systems Development* (pp. 1–18). Linköping Sweden: Linköping University.

Hollnagel, E., Woods, D. D., & Leveson, N. G. (2006). *Resilience engineering: Concepts and precepts.* Aldershot, UK: Ashgate.

Jones, R. A. (1986). *Emile Durkheim: An introduction to four major works*. Beverly Hills, CA: Sage.

Kohn, L. T., Corrigan, J. M., & Donaldson, M. (Eds.). (1999). *To err is human: Building a safer health system.* Washington, DC: Institute of Medicine.

Kowalsky, N. B., Masters, R. L., Stone, R. B., Babcock, G. L., & Rypka, E. W. (1974). *An analysis of pilot error related to aircraft accidents* (NASA CR-2444). Washington, DC: NASA.

Kuhn, T. (1962). *The structure of scientific revolutions.* Chicago, IL: University of Chicago Press.

National Safety Council. (2004). *Injury facts 2004 edition.* Itasca, IL: Author.

Orasanu, J. M. (2001). The role of risk assessment in flight safety: Strategies for enhancing pilot decision making. In *Proceedings of the 4th International Workshop on Human Error, Safety and Systems Development* (pp. 83–94). Linköping Sweden: Linköping University.

Reason, J. T. (1990). *Human error*. Cambridge, UK: Cambridge University Press.

Reason, J. T., & Hobbs, A. (2003). *Managing maintenance error: A practical guide*. Aldershot, UK: Ashgate.

Rochlin, G. I. (1993). Defining high-reliability organizations in practice: A taxonomic prolegomenon. In K. H. Roberts (Ed.), *New challenges to understanding organizations* (pp. 11–32). New York: Macmillan.

Rochlin, G. I. (1999). Safe operation as a social construct. *Ergonomics, 42,* 1549–1560.

Rouse, W. B., & Rouse, S. H. (1983). Analysis and classification of human error. *IEEE Transactions on Systems, Man, and Cybernetics, SMC-13,* 539–549.

Shappell, S. A., & Wiegmann, D. A. (2001). Applying Reason: The human factors analysis and classification system (HFACS). *Human Factors and Aerospace Safety, 1,* 59–86.

Slovic, P. (1992). Perception of risk: Reflections on the psychometric paradigm. In S. Krimsky & D. Golding (Eds.), *Social theories of risk* (pp. 117–152). Westport, CT: Praeger.

Vaughan, D. (1999). The dark side of organizations: Mistake, misconduct, and disaster. *Annual Review of Sociology, 25,* 271–305.

Woods, D. D., Johannesen, L. J., Cook, R. I., & Sarter, N. B. (1994). *Behind human error: Cognitive systems, computers and hindsight.* Dayton, OH: CSERIAC.

Sidney W. A. Dekker is a professor of human factors and system safety and director of research at the School of Aviation, Lund University, Sweden. He gained his Ph.D. in cognitive systems engineering from the Ohio State University in 1996.

Date received: May 8, 2005
Date accepted: January 17, 2006