SAND99-2868C

# Flawed Assumptions, Models and Decision Making: Misconceptions Concerning Human Elements in Complex System

**Chris Forsythe and Caren A. Wenner**
Statistics and Human Factors
Sandia National Laboratories
jcforsy@sandia.gov, cawenne@sandia.gov

## Introduction

The history of high consequence accidents is rich with events wherein the actions, or inaction, of humans was critical to the sequence of events preceding the accident. Moreover, it has been reported that human error may contribute to 80% of accidents, if not more (Dougherty & Fragola, 1988). Within the safety community, this reality is widely recognized and there is a substantially greater awareness of the human contribution to system safety today than has ever existed in the past.

Despite these facts, and some measurable reduction in accident rates, when accidents do occur, there is a common lament. "No matter how hard we try, we continue to have accidents." Accompanying this lament, there is often bewilderment expressed in statements such as, "There's no explanation for why he/she did what they did." It is believed that these statements are a symptom of inadequacies in how we think about humans and their role within technological systems. In particular, while there has never been a greater awareness of human factors, conceptual models of human involvement in engineered systems are often incomplete and in some cases, inaccurate.

## How Do We Think About Humans

The following sections discuss basic misconceptions that may consciously, but often unconsciously, shape perspectives on human involvement in engineered systems.

### Human as Machine

There are certain skills that greatly facilitate a career in engineering. One such skill is the ability to readily comprehend the interactions of different technical (i.e., mechanical and electrical) pieces of an engineered system. Most engineers are extremely adept at recognizing how one part interacts with the next and so forth for each part in a series. This same logic may be applied to human system components to identify the functions filled by humans and the consequences of simple human missteps. This is a valuable exercise that typically reveals numerous potential modes of failure. However, the ease with which potential human errors may be deduced through a systematic step-by-step analysis can create the illusion that the human system component is easily understood. mostly common sense, not unlike deducing the failure modes of a simple gear shaft or electrical circuit. Consequently, attention will often be focused on technical issues for which there is a thorough appreciation of the complexity of the problem, and well understood metrics (e.g., component reliability requirements).

The truth is that the human is typically the most complex and least understood system component susceptible to a more diverse range of failure modes than any other component of the system. Where engineers and analysts fall prey to the apparent efficacy of the mechanistic, or human as machine, perspective, the result can be an incomplete assessment that only captures the more obvious means by which humans may contribute to system failures.

### Random Behavior

On the surface, without an understanding of underlying psychological processes, a specific behavioral event may appear random. In tightening a series of twenty-eight screws, why was the tenth screw overlooked and not the fifteenth, or twentieth? After five years driving to work by the same route, why would one suddenly get in the wrong lane, activate their turn signal and make a wrong turn? These types of errors occur for no apparent reason and create an impression that they are unexplainable, a consequence of random behavioral processes. Two manifestations of this tendency to evoke randomness in explaining human behavior have been noted.

At a pedestrian level, there is a general willingness to accept that unusual behavior is a consequence of factors beyond comprehension. This is reflected in statements such as, "Who knows why anyone would ever make that mistake." At a theoretical level, the notion of randomness is inherent in certain applications of concepts from Reliability Engineering to human performance. In particular, repeated behavioral events have been represented using a statistical distribution with the

# DISCLAIMER

presumption that individual events are randomly distributed across that range.

There is a tremendous danger in evoking notions that behavior is the result of random processes. Behavior that is random cannot be predicted, nor controlled. Thus, where one willingly accepts randomness, one is helpless to do anything but accept the fallibility of humans. Oddly, this conclusion is often accompanied by a denial that vulnerabilities may exist (e.g., "that was an isolated event that will never happen again"), suggesting an unwillingness to acknowledge what one cannot explain.

Additionally, when rare occurrences of unintended behavior are attributed to random processes, there is also a tendency to cite "carelessness" or "inattentiveness" as the cause for the behavior. This simplifies the problem by allowing it to be isolated to a single individual. The solution becomes obvious, "Fix the person." Consequently, systemic problems may be masked and while the exact same behavior may never reoccur, other completely different unintended behaviors may follow.

### Isolation of Individual Human Agents

Engineered systems are typically complex and their analysis requires that to some degree, they be dissected into smaller units that are more amenable to detailed analysis. A product of this approach is the isolation of human system components with consideration of only those human-to-human interactions required by operational specifications (e.g., material transfers, proceduralized hand-offs). It has been observed that while complex operations may have formal channels by which human-to-human interactions are expected to occur, these formal channels are almost always accompanied by informal channels that are essential to day-to-day operations. This is often purely a product of human adaptation. Where there are inefficiencies and malfunctions in the system, people find ways to protect the system and keep it operating.

While a useful simplification, the assumption that individuals operate in isolation, with the exception of required interactions, may conceal complexity critical to the safety of the system. It may also cause the neglect of vulnerabilities resulting from unanticipated human interactions. Complex systems theories state that the number of states that a system may enter is proportionate to the number of entities and the number of interactions between those entities (Kauffman, 1995). Simplifications that dismiss informal human interactions may lead to a failure to recognize potential system states with significant consequences for the safety of the system.

### Predictability of Individual Human Agents

Earlier, it was noted that behavioral variations are often attributed to random variability and that this attribution may prompt beliefs that one is helpless to do anything but accept human fallibility. To some degree, this belief is predicated by an underestimation of the variability inherent to human behavior. When variability is believed to be uncommon, there is a tendency to attribute rare instances of high variability to random events.

In general, human variability complicates engineering design. Consider the design of equipment to accommodate people with physical dimensions ranging from the 5th to 95th percentile. This can be a difficult task and in this case, the variability may be quantified using parameters that are readily measured, unlike much behavioral variability. Whether done consciously or unconsciously, it is only natural to overestimate the constancy in human behavior. In fact, it may sometimes be necessary to the tractability of complex engineering problems.

The overestimation of constancy is reflected in a tendency to assume people are all the same or a given individual is always the same. This certainly simplifies systems design and analysis. However, the consequence is an engineered system that cannot tolerate the full range of variability humans may introduce to that system.

### Just So World

This misconception is closely related to the overestimation of human constancy. However, in this case, there is a tendency to assume that the system will operate as designed. This means that workers will follow written procedure and the written procedures are correct, engineered controls will be in place and operate effectively, training will accomplish the desired outcome, personnel selection will weed out the "bad apples," and in general, people will behave rationally and in the interest of the organization. Again, this presumption reflects a simplification of the engineered system that may be beneficial in allowing attention to be focused squarely on technical components of the system. However, applied exclusively without ample consideration of human factors, system complexity is ignored and critical failure modes may not be recognized.

### Erroneous Attribution of Mental Models

With software user interfaces, it is often observed that the designer, having intimate knowledge of the software, will mistakenly assume that the design is highly intuitive. Because the necessary sequence of operations seems obvious to the designer, they assume it will be obvious to everyone else. The designer will have a well-developed mental model with rich representations of the various features and how they interact. With a less developed mental model, there may be confusion leading to inappropriate behavior. The same applies to hardware interfaces, as well as written procedures. Furthermore, management decisions are often communicated with an implicit assumption that the underlying rationale is generally understood, when in fact, there are discrepancies that prevent the decision from being executed as intended. Furthermore, the mental model held by a designer or manager may be grossly inaccurate. It may reflect basic misconceptions and a lack of in depth knowledge of how people really do their jobs and the organization really operates. Familiarity promotes an assumption that others share the same mental model lessening the ability to recognize potential variability in the behavior of others, as well as flaws in one's own mental model. Consequently, there is often dismay when behavior is observed that is illogical in light of the assumed mental model.

### Disproportionate Attention to Line Workers

In conducting safety analysis, there is often intense scrutiny directed at the activities of line workers. "Line worker" is used generically to refer to those individuals with hands-on responsibility in hazardous operations. This attention is justified given that when a critical error occurs at the level of the line worker, there may be little or no opportunity for recovery. Those farther upstream than the line worker (e.g., designers, software developers, managers) typically enjoy the luxury of having multiple opportunities to detect and recover from critical errors and the presence of other factors that attenuate the consequences of their errors. Nonetheless, recovery mechanisms often fail and as will be illustrated by data presented in the following section, accidents may often be attributed to errors by individuals farther upstream than the line worker. Thus, inadequate attention to errors in engineering design, fabrication and assembly, management and other upstream functions may result in critical failure modes being overlooked.

## What Can Be Said About Human Error

Several studies have assessed various accidents to identify patterns associated with human error (Jarvinen & Karwowski, 1995; Lawrence, 1974; Salminet & Tallberg, 1996; Wagenaar & Groeneweg, 1987). The present paper discusses two relatively small-scale analyses that assessed the proportion of accidents attributable to human errors by different job functions.

### Initial Study

This study used a sample of forty incidents from the *United States Department of Energy Operating Experience Weekly.* Typical incidents involved electrocution, contamination, hazardous spills, etc. None of the incidents studied would be categorized as a high consequence event. However each incident is representative of the typical events in a chain of events leading to a high consequence accident.

Based on the incident reports, the immediate cause of each event was determined. Each immediate cause was assigned to one of the following job functions: Line Worker; Engineering Design; Maintenance/Fabrication; Transportation; Management; or Mechanical or Electrical Failure not directly attributable to human error. Figure 1 illustrates the proportion of accidents assigned to each function. It may be observed that while line worker errors accounted for almost a third of the incidents, errors in engineering design accounted for almost a quarter of the incidents. Furthermore, almost half of the incidents attributable to Line Workers involved errors using a non-intuitive user interface. It may be argued that these errors are also attributable to failures in engineering design. It should also be noted that there were several incidents attributable to errors during maintenance and fabrication, and errors by management.

### Follow-Up Study

As a follow-up to the initial study, a second sample of 47 incidents was obtained from the *DOE Operating Experience Weekly*. However, in the follow-up study, all errors contributing to each incident were identified, as opposed to only the immediate cause of the incident. Thus, several errors may be identified for a given incident. The Line Worker category was redefined to include fabrication, assembly, maintenance, transportation and other similar job functions. Additionally, the engineering category was divided into separate categories for Design, Software, Facility and Operations. As shown in Figure 2, when all errors are considered, the combined errors of each of the engineering disciplines exceeds that of the Line Workers, whereas the proportion of accidents involving management error remained the same.

As shown in Table 1, errors were further categorized with regard to error type. For this analysis, the engineering disciplines were combined. Due to the relatively small sample size, no attempt was made to quantitatively compare job functions.

Most importantly, it should be noted that very few errors were the simple omissions and commissions that would have been anticipated applying common analytic approaches. With little exception, errors involved relatively complex cognitive processes. Furthermore, a substantial proportion of the errors involved a failure to anticipate potential hazards. This occurred for all three job functions. It is believed that this failure is akin to a failure in hypothetical reasoning or stated differently, a failure to ask the appropriate "What if?" questions.

Figure 1. Percentage of Incidents in which the Immediate Cause was Attributable to Each Job Function

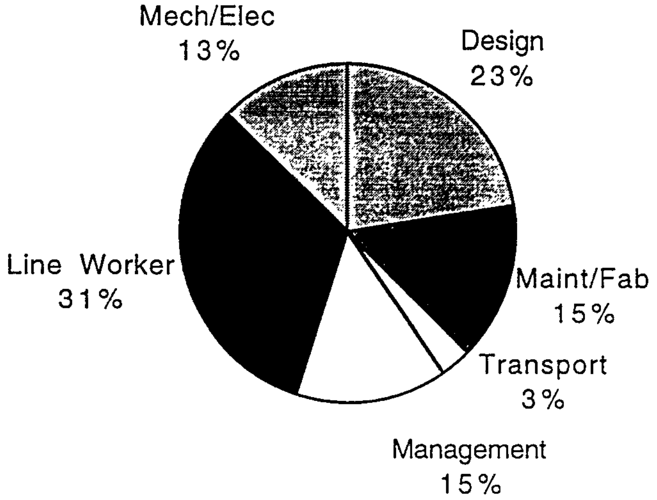

Figure 2. Percentage of Total Number of Incidents in which Errors Were Committed by Each Job Function

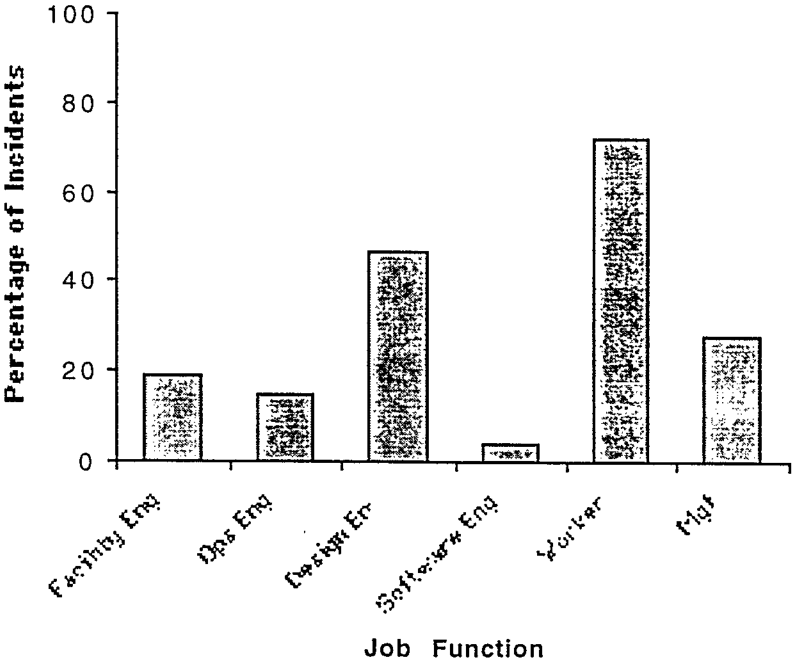

**Table 1. The Types of Errors Committed by Each Job Function**

| Error Type | Line Worker | Engineering | Management |
|---|---|---|---|
| Failed to Anticipate Hazard | | | |
| - Events | 9 | 11 | 6 |
| - Adequacy of Controls | 3 | 6 | 4 |
| - System Interaction | 6 | 9 | 5 |
| - Significance | 6 | 1 | 6 |
| - Human Error | 2 | 22 | 6 |
| | | | |
| General Cognitive Errors | | | |
| - Authorization | 3 | | |
| - Dynamic Properties | 4 | 3 | 1 |
| - Failure Discriminate | 1 | | |
| - Incomplete Mental Model | 3 | | |
| - Mistaken Mental Model | 7 | | 1 |
| - Placekeeping | 2 | | |
| - Functional Requirements | | 2 | |
| - Mismatch Reality and Artifact | | 1 | |
| - Material Properties | | 2 | |
| - Response Latency | | 1 | |
| - | | | |
| Failure Anticipate Requirements | | | |
| - Communicate Hazard | 1 | 2 | 1 |
| - Environment | | 2 | |
| - Functions | | 4 | |
| - Maintenance | | 9 | |
| | | | |

## An Alternative Way to Think About Human Involvement in Engineered Systems

Human interaction with technical system components is an inherent feature of engineered systems. These interactions may be intended, part of the system design, or unintended.

To begin, consider interactions that are intended. For each interaction, there is variability associated with that interaction. When tightening a bolt to a specified torque value, the desired torque may be exceeded to the point where the threads are stripped or the bolt is broken. Similarly, an insufficient torque may be applied or the torqueing operation may be omitted altogether. As Illustrated in Figure 3, for each interaction between humans and technical system components, there is a range of variability for which the system is tolerant and beyond that range, the system is intolerant of variability and may be susceptible to undesired consequences.

The objective is to identify interactions that afford little tolerance for variability and assure variability does not exceed the limits for which the system is tolerant. This may be accomplished in two ways. First, measures may be taken that increase the tolerance for variability (e.g., independent verification, mechanical limits, lock-outs). Second, steps may be taken to reduce variability (e.g., written procedures, training, job aides). While numerous counterexamples may be cited, the first approach generally involves technical solutions, whereas the second approach usually emphasizes administrative solutions.

Humans may also interact with technical system components in ways that the system designers never intended, nor anticipated. Again, the result may be that humans introduce variability that exceeds the tolerance limits of the system. The same solutions discuused for intended interactions also apply for unintended interactions. However, it is much more difficult to anticipate the range of potential variability possible with unintended interactions.

The dilemma is one in which measures must be taken to assure systems do not experience variability that exceeds the tolerance of the system, but the variability cannot be anticipated from either a qualitative or quantitative perspective. A solution is proposed that seeks to lessen variability by identifying and addressing the organic properties of humans that represent the underlying factors responsible for the variability humans introduce in their interactions with engineered systems

Table 2 lists eight organic properties that contribute to variability in human behavior and describes the effect of each property on engineered systems. . Knowing these properties and how they contribute to variability in human behavior, controls may then be proposed that lessen this variability. For example, variability that arises due to the susceptibility of humans to fluctuations may be partially controlled through work schedules, monitoring, pacing or other related measures. This is only one example and further work is required to fully describe the means by which organic properties produce variability in human behavior. Particularly, biological limitations that constrain cognitive capabilities will require considerable attention since as Table 1 illustrates, actual incidents are typically the result of failures in reasoning and hypothetical thinking, not simplistic omissions and commissions.

## Conclusion

The previous sections have described misconceptions that contribute to the failure to anticipate interactions between humans and technical components of engineered systems that are critical to system safety. In many cases, these interactions may result in variability that exceeds the tolerance of the engineered system. Since the anticipation of all potential forms of variability is considered to be intractable, an approach is offered that focuses on controlling organic properties of humans that are the source of variability in human behavior.

## Acknowledgements

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

## References

Dougherty, E.M. & Fragola, J.R. (1988). *Human reliability Analysis*. New York, Wiley.

Jarvinen, J. & Karwowski, W. (1995). Analysis of self-reported accidents attributed to advanced manufacturing systems. *The International Journal of Human Factors in Manufacturing, 5(3)*, 251-266.

Kauffman, S. (1995). *At Home in the Universe: The Search for the Laws of Self-Organization and Complexity*, Oxford Press, Oxford.

Lawrence. A.C. (1974). Human Error as a cause of accidents in gold mining. *Journal of Safety Research, 6(2)*, 78-88.

Salminet, S. & Tallberg, T. (1996). Human errors in fatal and serious occupational accidents in Finland. *Ergonomics, 39(7)*, 980-988.

Wagenaar, W.A. & Groeneweg, J., (1987). Accidents at sea: Multiple causes and impossible consequences. *International Journal of Man-Machine Studies*, *27*, 587-598.

**Table 2. Organic Properties Introduced by Humans to Engineered Systems**

| Property | Manifestation |
|---|---|
| 1. Constant Fluctuations | Individual performance varies in accordance with biological rhythms |
| 2. Susceptible to Perturbations | Minor events can have a major impact on performance |
| 3. Adaptive | There is constant adaptation to actual and perceived threats, and resource allocations |
| 4. Instinctual | There are ever-present instincts that may precipitate behavior contrary to the rules of the system, and personal well-being |
| 5. Propagation | Behavior does not occur in isolation, but is propagated through observational learning and cultural transmission |
| 6. Self-Organizing | Entities/subsystems created that overlay engineered system |
| 7. Meta-Systems | Assumptions, beliefs and rules introduced that supersede engineered system (e.g., values, customs) |
| 8. Biological Limits | The capabilities of the engineered system are bounded by inherent physical and cognitive limitations of human participants |

**Figure 3. Engineered Systems Exhibit Differing Degrees of Tolerance for Human Variability**