Chapter 6: Error

Fallibility is intrinsic to human behavior. As Alexander Pope (1688 – 1744) has put it, "to err is human." That might explain why there are many terms (e.g., errors, mistakes, failures, blunders, faults, slips and lapses) to describe situations where things do not happen as we expect. Regardless of what we call them, human error is pervasive. Just like gravity and weather, it is an unavoidable fact of life that occurs every day and every hour – as we speak, as we drive, and as we do other daily routines.

The innate fallibilities of people are built into every system that involves human beings and makes those systems vulnerable to flawed human actions. As increased complexity becomes a crucial feature of contemporary human-technical systems involving mechanization, computerization and automation, it increases the prospects of human error and makes such failures increasingly dangerous. Statistics show that human errors are implicated in over 90% of failures in the nuclear industry, over 80% of adverse events in healthcare systems, over 80% of failures in the chemical and petro-chemical industries, over 75% of marine casualties, and over 70% of aviation accidents. The potential devastating consequences of erroneous human actions are exemplified by severe disasters such as Chernobyl, Bhopal, Challenger, and Air Florida Flight 737. All these names are etched in our consciousness for substantial property damage, and injuries, as well as deaths that have occurred over extended periods of time and across large geographic space, and have received extensive public, political, and scholarly attention. Human contributions to economic loss were illustrated in a study conducted by The National Institute of

Standards and Technology (NIST) (Tassey, 2002) that considered major US software development firms. It was reported that software engineers spend an average of 70 - 80% of their time testing and debugging, and it takes them an average of 17.4 hours to fix a bug, costing the US economy over $50 billion annually.

Error from the brain's perspective

From the perspective of the brain, you are right all the way up until the point that you realize that you are wrong. This expression captures the common experience that accompanies our committing errors during everyday activities. We sit the kettle on the stove, turn on the burner and walk away waiting for our water to boil. Everything we have done seems fine and we continue with our morning routine for fixing breakfast. From the brain's perspective, everything we have done is correct, we have no reasons for concern and we are making progress toward our goal to have breakfast and get ready to start our day. Then, for some reason, we realize that we forgot to put water in the kettle. In a momentary realization, everything we had previously taken for granted in carrying out our morning routine is called into question as we berate ourselves for being so absent-minded and putting our family and home in danger. This is our common experience of having committed an error. Everything is fine all the way up until the point that it becomes apparent that everything is not fine.

Within experimental research studies, this common experience of realizing that one has committed an error has been studied and the underlying brain mechanisms elucidated. A typical experimental paradigm presents a test participant with two buttons and an indicator that informs the subject which of the two buttons to press. The objective is to press the correct button as fast as possible once the indicator appears. In this paradigm, known as a choice-reaction time paradigm, subjects usually commit numerous errors as they attempt to maximize the speed of their response. On those trials in which the subject commits an error, there is a distinct neurophysiological signal that accompanies this realization (Holroyd & Coles, 2002). Specifically, there is a wave of activity measurable using EEG that emerges at the point in which the motor response commences, peaking about 80 msec afterward, and spreading across a broad swath of the brain (i.e. error-related negativity). This signal has been attributed to a general-purpose system within the brain that functions to detect the occurrence of errors and prompt a re-orientation of attention in response to an error. The system is differentially responsive to motivational considerations with it demonstrated that when rewards are provided for successful performance, the magnitude of the signal varies in accordance with the corresponding payoff (Gehring et al, 1993). Furthermore, the magnitude of the error signal varies in response to the magnitude of the error. For example, in a four-choice reaction time task where subjects responded using both hands and two different fingers on each hand, the magnitude of the error-related response was larger the more errors diverged from the correct response (Bernstein, Scheffers & Coles, 1995). For example, there would be a larger amplitude response when the subject used the wrong finger on the wrong hand, as opposed to merely responding with the wrong finger or the wrong hand.

The realization by the brain that an error has occurred provides the basis for the brain's response to errors. A study was conducted in which subjects were presented an indicator light and asked to wait one second before they pressed a button (Miltner, Braun & Coles, 1997). Following the button press, the subjects were given no feedback concerning their performance. Then, at the end of a trial, the subject was told whether their performance on the trial had fallen within the criteria. An error-related response was seen when the subject received the feedback suggesting that it was the knowledge of how well they performed, which had been dissociated from the actual performance, that induced the brain response. The error response emanates from a region of the brain known as the anterior cingulate cortex (Dehaene, Posner & Tucker, 1994). It has been suggested that this region serves as a comparator. For a given response, there is an intended outcome and the anterior cingulate cortex compares the actual outcome to the expected outcome and when there is a discrepancy, emits a response that orients attention to the activity. Furthermore, Holroyd and Coles (2002) have proposed that there is a link between the anterior cingulate cortex and the reward circuits of the brain such that the comparative function of the anterior cingulate cortex serves to modulate the level of reward (i.e. dopamine release in association with the response). Thus, through a continuous comparison of actual and intended outcomes by a generic and highly flexible response monitoring circuit, the error response of the brain serves to continually refine behavioral performance.

Organizational approach to human error

Since human fallibility is inevitable, within the context of an organization, it is counterproductive to attempt to completely eliminate errors or establish goals such as flawless performance or perfection, which is often advocated by engineering system designers and/or management. Rather, we first need to adopt a perspective that recognizes the existence and potential occurrence of human errors, and then focus on goals of developing means and measures to prevent errors from occurring and minimize their consequences when they do occur. This approach mirrors the functional organization of the brain where a premium is placed upon the detection of errors and error monitoring provides a basis for continual refinement of performance. Within an organization, such an effort is needed continuously. As emphasized by Weick and Sutcliffe (2001), in contrast to the pervasive nature of error, what is not pervasive are well-developed systems, processes, and skills to detect and contain errors at their early stages. Echoing this perspective, there is a long tradition of research on human factors, safety management and risk analysis that aims to develop approaches for assessing system reliability, evaluating potential risks of specific errors, identifying error causes, and preventing errors.

Given their potential costs and risks (adverse effects), we normally ascribe errors a negative role in our life with a deeply ingrained aversion. Compared to the extensive efforts devoted to preventing errors and evaluating their risks, the potential benefits of errors have been largely overlooked. The history of human civilization is full of examples of innovations and discoveries inspired by errors. As James Joyce noted poetically, "mistakes are our portals of discovery." By allowing for variation far beyond what was expected, errors can provide brilliant insights to stimulate creativity and invite us to explore new paths to "success on the far side of failure" (Thomas J. Watson, Sr., 1874-1956). We owe thanks to errors for many of the biggest game-changing inventions and discoveries of our times, including the lifesaving antibiotic penicillin, smallpox vaccine, pacemakers, dynamite, plastic, radioactivity and many others. These inventions and discoveries not only have fundamentally changed every aspect of our life, but have reshaped major industries or created entirely new ones. Furthermore, as a natural part of learning and training, errors indicate performance deficiencies and can, under some circumstances, provide informative feedback about where knowledge and skills need further improvement (Keith, 2010). Therefore, understanding how to learn from errors, as naturally occurs with the brain, is as important as preventing errors. This is not to say that we should forego the prevention of errors. Rather, we should foster a positive attitude toward errors when they occur by recognizing their potential instructive value in germinating new insights, particularly if the outcome of the error is an incident rather than an accident. Scott Adams expressed this sentiment in his famous line, "Creativity is allowing yourself to make mistakes. Art is knowing which ones to keep." Errors will cause losses, but they are like bitter medicine, of which we need to take a few gulps to spark the creative genius inside of us.

Although human error is a concept rooted in the discipline of psychology that describes an aspect or phenomenon of human behavior, it has increasingly drawn attention and interest from a wide range of domains outside of psychology, including engineering, medicine, social science and many others. This is not a coincidence but a consequence of its paramount importance for the safety of socio-technical systems and our desire to avoid disasters in the real world. Particularly, the past several decades have seen a multitude of developments made through cross-disciplinary efforts, with an impetus to understand the human contribution to safety and risk in complex systems from individual, organizational, social, and design perspectives. However, human error is not a simple topic. Despite extensive research efforts and significant achievements, progress in the study of error is still slow.

Confusion regarding the term "human error"

The term of "human error" has been used in our daily language throughout history as if the term were well understood and defined. On the contrary, although many working definitions have been proposed, there is an agreement that the term is difficult to define, with there being little consensus among the definitions. It is not the intention of this chapter to provide an ideal definition for the term. Here, the term is used to refer to an activity that (1) was not intended by the actor; (2) is undesirable with respect to relevant rules or an external observer; or (3) resulted in a task or system diverging from its acceptable limits (Senders & Moray, 1991).

The terminological confusion about human error largely arises from the fact that the term is loosely used with a great diversity of notions and interpretations. As noted by Hollnagel and Amalberti (2001), while the term has a connotation of something that should be avoided, it is used with many denotations, each of which can be misleading in understanding the nature of error. In this section, we will discuss the controversies and ambiguities about the meaning of the term.

Human error as judgment

To understand the term, "human error," it is important for us to realize that when we say that a human error occurred, we are making a judgment by comparing, either explicitly or implicitly, the so-called error to some performance standard or criterion. That is, without a comparison, it is not meaningful to say that an error has been made even though the outcome is undesirable; an error can only be said to exist if there is a clear performance standard to define the criterion for an acceptable response or outcome.

The judgment regarding error implies that human performance is either correct or incorrect – a binary distinction. Under some circumstances, the distinction can be made rather easily and objectively with a clear performance criterion. For example, for tasks that require a discrete action, such as turning on a pump, the objective correctness of the action can be easily specified. However, the binary distinction between right and wrong is not always justifiable for human performance. For instance, when there are multiple alternative strategies to achieve a goal, even though some strategies may be less optimal and lead to an imperfect outcome, they cannot be considered error as long as the outcome remains within the specific tolerances required for achieving the goal. Furthermore, even for the optimal strategy, the outcome may vary as a product of the natural and irreducible variability in human performance. Under these circumstances, it is difficult to define what is meant by incorrect performance except for situations where human actions are completely inappropriate for the purpose or beyond acceptable limits of a specific task. That is, the defining qualities of the judgment may be unclear and thus, may be prone to interpretation, especially when the criterion involves human internal, unobservable cognitive constructs such as intentions and purposes. Hence, the term "human error," to some extent, involves relativism and subjectivity. An error in one person's eyes may not be considered an error by somebody else, depending on their perspectives. There is a range of possibilities from the entirely correct to the totally incorrect, and the distinction between right and wrong is normally a matter of degree rather than absolutes (Hollnagel, 2007).

An interesting example illustrating the subjectivity in judgment regarding human error is presented by Hollnagel and Amalberti (2001). In a study analyzing Air Traffic Controller (ATCO) errors, two observers watched a controller's training performance on a full-scale simulator and took note of all the errors that they observed. Although the observers witnessed the same performance, there were substantial differences in the number of errors that they noted.

More importantly, there were only a small number of errors noted by both observers.

From the brain's perspective, the subjectivity in judgment regarding human error is evident in the physiological response accompanying erroneous performance. For example, it has been demonstrated that there are multiple components of the error-related response of the brain. There is an initial component that has a negative polarity that is present whether or not an individual is aware of having made an error (Nieuwenhuis et al, 2001). In contrast, there is a second later component with a positive polarity that is generally absent in conditions where a subject commits an error, but is unaware of having done so. When subjects are asked to evaluate their own performance and estimate the magnitude of their errors, there is a correlation between the magnitude of the error response and the individual's subjective appraisal regarding the extent to which they have erred (Scheffers & Coles, 2000). Furthermore, when faced with an erroneous outcome, the brain responds differently depending on whether an individual perceives that they are the source or cause of the error, as opposed to their believing that they had no role in causing the error (Knoblich & Sebanz, 2005). These findings point to the subjectivity inherent in human error. From the brain's perspective, error is not absolute, and one's response to error differs in response to whether one perceives that they are the cause of the error.

Human error as cause

In our daily language, we often say that an undesirable event occurred because of human error. In this context, the term "human error" is used with a denotation of being the cause of the event. Logically and philosophically, such a denotation represents a reverse causality (i.e., reasoning from effect to cause) and implies that human error is an inference, judgment or rationalization made through reverse causation. Hence, the use of the term, "human error," in this sense risks falsely associating an observed effect with a presumed cause (Hollnagel and Amalberti, 2001).

Human error may be one of the multiple possible causes for observed performance or one variable in the complex process leading to the performance. Therefore, the denotation is, to some extent, misleading in that it attributes all the possible causes and associated complexity to an over-simplified plausible explanation labeled "human error." This attribution is normally made when search for the cause of performance is carried just far enough so that there is insufficient evidence to "blame" technological systems, such that human error constitutes a catch-all satisfactory explanation (Hollnagel and Amalberti, 2001; Woods et al., 2010). Such an explanation is often convenient, as well as sufficient, in accident investigations, because, as Perrow (1986) noted, "Finding that faulty designs were responsible would entail enormous shutdown and retrofitting costs; finding that management was responsible would threaten those in charge, but finding that operators were responsible preserves the system, with some soporific injunctions about better training" (p. 146).

Since the denotation suggests that a meaningful cause has been identified, namely the human, it stifles the search for a systematic explanation of why the "unsuccessful" performance occurred and paints a partial picture of the undesirable event without considering other potential factors (see previous discussion regarding contextual factors). The misleading implications of this denotation are that: (1) error is simply associated with or originated with humans; (2) humans are responsible for the error; and (3) error prevention can be achieved by changing humans or reducing their role in complex systems (Woods et al., 2010). As pointed out by Woods et al.

(2010, p. 4), error is a symptom rather than a cause and should serve as a call for further probing and investigation. The implications may create confusion or ambiguity in distinguishing between the *cause* and *agent* of error. As discussed previously, a human operator may be the agent but not the cause of an error, because the decision or action of the human operator may seem erroneous in the sense that it leads to undesirable results, but the decision or action can be sensible in the light of its context.

Human error as consequence

When we conclude that an error was caused by human action, the term "human error" is used to characterize the observable outcome of the action. Such a denotation is commonly used in the probabilistic risk analysis (PRA) community to refer to human-caused failures of a system or function. One problem with this detonation is that it does not provide much insight concerning error except stating that the planned action fails to achieve its desired goal.

Another problem is that judgments concerning error, in this case the consequence of an action, can only be made after the error has occurred. Such judgments represent after-the-fact inferences based upon observations of unsatisfactory results, and consequently, are prone to hindsight bias. As Woods et al. (2010) pointed out, with the benefit of hindsight, we know what factors were critical to the results and tend to assume that knowledge was available before the results were known. As a result, we are biased to oversimplify or trivialize situations, as well as the mechanisms that could have affected the results, exaggerating what could have been known in foresight.

Human error as action, event, or process

When we say an error has occurred, or an event or process has gone wrong, we address the manifestation of the error without consideration of the cause. Furthermore, we do not address the consequence, such as how the error may be manifested within overt action and how it can be observed in an awry process. One problem with this denotation is that it, to some extent, implies that human error is intentional, an observable human function, or an observable category of human performance. As discussed above, human error is inferred from observations and involves judgments made in hindsight (Hollnagel, 1983; Woods et al., 1994).

There is no good reason to suppose a close coupling between action, event or process and the subsequent outcome or consequence (Woods et al., 1994). A slip may cause momentary embarrassment and trifling inconvenience in a forgiving circumstance, but could lead to a catastrophic accident in a safety-critical system such as a nuclear power plant. As Reason and Mycielska (1982) put it, the difference lies not in the nature of the action, event or process, but the extent to which its circumstances will produce penalties. In addition, an "incorrect" action or an imperfect process does not necessarily lead to failures, and actions or processes can be corrected before they generate an undesirable outcome or consequence. On the other hand, even if there are good reasons for performing some actions, they may still lead to unwanted consequences.

Interactive nature and complexity of human error

One important notion that is missing in the denotations or interpretations discussed in the previous section is the interactive nature and complexity of the processes that produce human performance. Error is the product of the dynamic interaction between human, and system, or task, to achieve a specific goal in a specified context. This section discusses three important elements involved in the processes underlying error: *intention*, *cognitive and neuropysiological mechanisms*, and *context*.

Intention

Intention includes the motivation to interact with a system or perform a task, and the goal(s) of the interaction. The notions of error and intention are inseparable (Reason, 1990, P5). Error is not a human function. We do not *commit* an error intentionally (i.e. error is not the result of an intention); however, error is only applicable to intentional acts where there was *prior intention*, meaning the intention was formed prior to performing the act (Searle, 1980). Note that some intentional actions, such as spontaneous behavior, are carried out without plans. In such cases, no mismatch between achieved results and the goal is present to evaluate whether a specific action is correct or not (i.e., whether the action deviates from intention). Hence, it is only meaningful to talk of error when we act on a prior intention.

According to Reason (Reason, 1990, P5), prior intention is comprised of two elements: (1) the end-state to be attained, and (2) the means to achieve the end-state. Based on the two elements, error can be classified into two forms (Norman, 1983:

- *Slips* (or *lapses*) are unintended actions (i.e., actions that deviate from prior intention) that do not achieve the intended end-state. They represent a discrepancy in executing intended actions
- Mistakes are intended actions that do not achieve the intended end-state because of the mismatch between the prior intention and the intended consequences. Note that when

there is a mismatch but we do not act on that intention, the mismatch may sometimes be called a mistake but will not result in an unwanted outcome.

Discussion of intent generally presumes some degree of conscious awareness. However, in research summarized in Chapter 2, it has been demonstrated that activity is present within the brain that is predictive of the intent to act and the specific nature of the action several seconds prior to the subject becoming consciously aware of their intent to act (Soon et al, 2008). Furthermore, conditions may be distinguished in which conscious awareness is focused on the intent to execute a specific action, as opposed to consciously focusing on the actual execution of the action, based on activity occurring in a region of the brain known as the pre-Supplementary Motor Area (Lau et al 2004). In a typical experimental paradigm, subjects are asked to choose between one or more actions and then hold this intent in mind for some period of time until cued to execute their response. Based on activity observed within the brain, the specific choice may be reliably predicted. Once the action begins, activity shifts to other regions of the brain as the subject executes the action and monitors their performance (Haynes et al., 2007). These brain processes underlying the intention and execution of an action are in play, whether or not one is consciously aware of their intentions or consciously attends to the action. Notably, conscious awareness is not essential, and an intent may exists within the brain to carry out a specific act, without conscious awareness of this intention. These observations confuse our common conceptions of intentionality and error. Conscious awareness of the act, including the erroneous nature of the act, may not occur until the act is either underway, or an individual has reached a point-of-no-return, with the erroneous act being unstoppable.

Cognitive and neurophysiological mechanisms

Information-processing model

Consideration of human error from a cognitive perspective has largely emphasized information-processing theory, whereby the human brain is conceptualized as an information processer, which processes and stores information much like a computer. This theory has had a profound impact on many disciplines and applied areas, such as human-computer interaction (HCI), human factors and ergonomics, and human reliability analysis (HRA). Figure 1 illustrates a typical information-processing model. Human cognitive processes are represented as serial or linear stages, with there being different memory stores (e.g., short-term memory and long-term memory) and processors. Similar to a computer, information is transferred from one store or processor to another. For example, consider an operator responsible for monitoring a complex process. A dynamic mental model of the actual process under supervision is developed through training and stored in the operator's long-term memory. Actual process information is first received through the operator's sensory receptors (in this case mainly visual and auditory receptors). Next it is interpreted and organized by the operator's perceptual processors to update the operator's mental model. The synchronization takes place continually at a subconscious level during process monitoring. When there is a disagreement between the actual process and the mental model, the operator's attention will be focused on the discrepancies, and conscious cognitive activities will take over to select an action to respond based on knowledge from the operator's long-term memory. The selected action will be executed by the operator's motor processors. Finally, the outcome of the operator's action will be registered through sensory processes creating a feedback loop to initiate a new information-processing cycle (Ivergard & Hunt, 2008).

Insert Figure 1 about here

An implication of the information-processing model is that an error is the result of subpar performance of human cognitive and motor functions of the information-processing cycle. Furthermore, the model assumes that a primary root cause of undesirable performance is the limited capacity or inherent “design defects” of the human information processors (Hollnagel, 2007). As shown in Figure 1, the whole information-processing cycle can be divided into two stages: *cognitive information processing*, which consists of information reception, identification, interpretation, and decision making, and *manual task execution actions*. Corresponding to this division, human error can be separated into two basic categories: *planning failures* and *execution failures*. This categorization corresponds to the aforementioned slips/mistakes categorization links errors to specific levels of cognitive processing. Planning failures or mistakes occur in the cognitive information processing stage, while execution failures or slips (lapses) arise in the manual task execution stage.

Neurophysiological mechanisms

While information processing models have been extremely influential, their depictions of brain processes are questionable. At best, the functioning of the human brain is weakly analogous to the functions of a digital computer. Consequently, adherence to an information processing model may prompt erroneous conclusions regarding the basis of human error. A consideration of the key differences between the brain and a digital computer elucidate several sources of human error within brain functions.

Whereas computers involve digital information processing, the brain is largely analog, with its operations being continuous and often, non-linear. The rate of firing and the synchronization of firing across neural circuits are essential to information processing within the brain. Consequently, performance can be variable depending on the relative timing of events. For example, visual perception is characterized by synchronous activity in the 8-13 Hz (alpha) bandwidth. It has been shown that the likelihood of detection of a visual stimulus varies in accordance with the phase of these cycles such that a stimulus presented during the ascending portion of the cycle is more likely to be detected than a stimulus presented during the descending phase (Busch, Dubois & Van Rullen, 2009). Similarly, variability in trial-to-trial performance has been linked to spontaneous fluctuations in the oscillatory activity of neural circuits (Fox et al, 2005). Thus, it is misleading to think of the brain as a digital processor carrying out a sequence of instructions. Instead, the brain is more appropriately conceptualized as a collection of oscillating neural assemblies capable of processing information through the phase and harmonic synchronization of oscillatory processes (Singer, 1993). Accordingly, human error may arise as a product of the inherent variability in these processes, including factors that accentuate the variability inherent to these processes (e.g. fatigue, age) or as a result of perturbations affecting these processes (e.g. emotional responses).

With a computer, memory is accessed through reference to a precise address corresponding to its location on the memory drive mechanism (i.e. byte-addressable memory). This design allows for great precision in recalling specific memory records. In contrast, memory within the brain is "content-addressable," meaning that recall is based on reference to the content of the memory. Thus, when recalling a specific memory, various cues to the memory contents serve as the basis for accessing and activating patterns of neural activation corresponding to the memory. Furthermore, within the brain, memory is subject to spreading activation. Through spreading activation, recalling a given item from memory primes other related items, and as a result of this priming, recall of related items is facilitated, with related items often consciously recalled.

Compared to a digital computer, memory processes within the brain are imprecise and subject to interference. Consequently, errors may arise due to erroneous recall. For example, with routine activities that have become highly automated, an irrelevant cue is often sufficient to trigger activation of an incorrect sequence of actions. This can be seen during daily routines when you lose track of where you are in the routine and repeat a step (e.g. brush your teeth twice or put the milk away before pouring some on your cereal). Likewise, interference attributable to spreading activation can impair accurate memory recall. For example, when trying to recall the name of someone that you have not seen in a long time, it is easy to mistakenly retrieve the wrong name based on the names of other people that you had known about the same time being more accessible within memory. In this situation, memory is accessed on the basis of the context in which you had known the person, but there were other people associated with this same context and in recalling the context, the wrong name was triggered. Yet, while the memory mechanisms of the brain are prone to error, the brain is often capable of retrieving memories given scant cues, or at the least, recalling a partial memory. In contrast, computers are generally incapable of retrieving items from memory without specific reference to the memory address, with memory retrieval being all or none.

Whereas computers are modular and carry out serial operations, the brain operates in a manner that is massively parallel. Previous chapters have discussed the manner in which the brain operates in a parallel fashion. During everyday experiences, the entire brain is active in processing and integrating sensory inputs, interpreting events and coordinating responses. This includes processes occurring at both conscious and unconscious levels. All of these processes are then interwoven to create our memory for these experiences. Consequently, our activities are context-sensitive. This means that given appropriate contextual cues, memories may be retrieved that would otherwise be inaccessible. However, it also means that placed in the wrong context, memory retrieval may fail or inappropriate memories may be retrieved. Likewise, misinterpretation of the context can lead to retrieval of inappropriate routines. For example, an examination of accidents often reveals that the operators misinterpreted the situation and given their interpretation of the situation, the seemingly erroneous actions were actually quite appropriate. The parallel nature of the brain lends itself to operational processes that are contextually-based meaning that the brain creates an integrated representation of events. Then, as one moves between contexts, the brain shifts from one contextually-linked collection of memory representations to another. This implies that for effective performance, there must be an accurate interpretation of contexts, with the failure to accurately interpret contexts leading to retrieval of erroneous and inappropriate memory representations.

As a result of the parallel manner in which the brain operates, the operations of the brain tend to be non-discrete. In contrast, with a computer, operations are serial and discrete. Consequently, with the brain, there is potential for interference between related operations. When there is activation of a given operation, this activation produces spreading activation to other related operations. For example, while fixing breakfast, the act of pouring a bowl of cereal produces spreading activation to the act of pouring milk onto the cereal. As a result, if the next step is to take orange juice from the refrigerator, it is not unlikely that will erroneously pour the orange juice onto the cereal. While specific steps may be carried out one at a time, spreading activation and the resulting priming of memory representations, with multiple representations being simultaneously activated, creates an operational environment where individual actions overlap with one another and can interfere with one another and occasionally, actions may be performed out-of-sequence.

Whereas computers have a system clock and operations are precisely timed relative to this clock, the timing of brain processes is variable. The extent of this variability can differ with there being associated effects on performance. For example, as one gains proficiency with a motor task, there is a reduction in the variability of neural processes involved in performing the task. This was demonstrated in a go/no-go paradigm in which subjects were presented an initial cue to warn them that the cue prompting their response was impending (Churchland et al., 2006). During the interval, there is activation of the premotor regions of the cortex that corresponds to the preparation to respond. With practice, there is reduced variability in neural activation during the interval and with this reduced variability, there is a reduction in response time. In other research, it has been shown that variability in neural responses impact perceptual experience. A study was conducted using the Rubin vase-faces picture for which subjects viewing the figure from one perspective report seeing a vase and from an alternative perspective, report there being two faces (Hesselmann et al, 2008). The researchers noted that there were ongoing fluctuations in neural activity and when presentation of the figure corresponded to elevated activity in the fusiform gyrus, which is a region of the brain associated with recognizing faces, there was a greater tendency to report seeing two faces, as opposed to seeing a vase. This finding suggests that our propensity to interpret sensory information one way or another is a product of the relative timing of ongoing fluctuations of neural activity. The operation of the brain involves the coordination, or synchronization, of neural activity, both at a local level corresponding to specific functional circuits, as well as at distributed levels for production of a response that integrates the activity of various functional circuits. The capacity to achieve synchronization of neural circuits can vary with the result being variability in behavioral performance. Furthermore, individual differences in intrinsic levels of neural variability, and corresponding cognitive abilities have been linked to the level of connectivity within the prefrontal cortex (Ullen et al, 2008). Thus, there appears to be a neural substrate underlying individual differences in the ability to generate coordinated neural activity, with associated impacts on cognitive performance.

Macrocognition

From a psychological perspective, the information-processing model in Figure 1 has proven inadequate to describe the true complexity and dynamics of human cognition. Human cognition involves a great deal of simultaneous, parallel, and circular processing. In addition, the information-processing model cannot adequately account for creativity, illogical thinking and sparks of genius (Anderson, 2000).

Recognizing the limitations of the information-processing theory, psychological research has explored alternative theories to elucidate the complexity of human cognition. Among these theoretical developments, the theory of macrocognition has become imminently useful. Macrocognition is a term originally coined by Cacciabue and Hollnagel (1995) to describe cognition in real-world settings. It focuses on the nature of human performance in "the field," where decisions often are very complex and must be made quickly, by domain experts, in risky or high-stakes situations (Klein et al., 2003). In contrast, microcognition is typically the focus of tightly controlled laboratory research, which aims to elucidate the building blocks that underlie complex cognition. There are a large number of microcognitive models focused on different aspects of human cognition.

According to macrocognition theory, a marcocognitive function is the high-level mental activities that must be successfully accomplished to perform a task or achieve a goal in a naturalistic environment (Letsky, 2007). Although there are a number of different macrocognition models, none of them were developed as a generic model applicable across different work domains. However, there is a general consensus among various models that there are five macrocognitive functions: *detecting and noticing*, *understanding and sensemaking*, *decision making*, *action*, and *team coordination* (U.S. NRC, 2012; Whaley et al., 2012). Whaley et al. (2012) adapted the macrocognition concept to operator behavior in nuclear power plants. They identified cognitive mechanisms corresponding to the processes by which cognition takes place in the work environment. The operation of these mechanisms is considered crucial to successful performance. If part of the process fails, (either internal or external to the human,) the failure may manifest itself as a macrocognitive function failure.

Over 300 mechanisms have been identified along with the associated boundary conditions that can lead to the failure of the macrocognitive functions. The mechanisms are then clustered into categories (i.e. *proximate* causes) based on the effects of the failures. A proximate cause is the result or manifestation of the failure of a mechanism and thus, can be readily identified as the basic contributing cause of the failure. For example, one cognitive mechanism states that, "important sensory cues must be sufficiently salient to be easily detected by higher cognitive functions." Failure of this mechanisms results in the cue not being perceived. Thus, "cue/info not perceived" is an identifiable cause (i.e. a proximate cause) for failing to detect a cue (the detecting and noticing macrocognitive function). The proximate causes for the failure of the five macrocognitive functions are listed in Table 1.

Insert Table 1 about here

Through establishing the direct linkages (i.e., causal relationships) of cognitive mechanisms to observed outcomes of performance, the macrocognitive model provides an analytical rather than empirical approach to address different classes of errors (e.g., errors of commission and errors of omission), using the same set of cognitive mechanisms. Furthermore, since the model addresses failures at the level of cognitive mechanisms, which is the basic fabric of human performance in different situations, the model can be readily extended to various domains.

Context

All human actions are performed within a specific context and are subject to a complex range of situational or environmental influences that bear on our ability to accomplish a task, such as training, procedures, human-system interface, communication standards, and so forth. The relevance and importance of contextual elements may vary with the situation, with contextual elements often being difficult to identify and measure.

Retrospective analyses of actual operational events in industrial complexes (e.g., nuclear power plants) show that human operators generally perform routine tasks well with reasonable, natural variability. Human performance problems often arise in unusual circumstances, where operators perform actions that are not required and, during an accident, may worsen conditions. An isolated analysis of individual operational events may leave an impression that failures are caused by operators' illogical actions under system- or event-specific circumstances. However, a deeper scrutiny from a holistic perspective suggests this impression is a biased and simplified explanation for the undesired human performance. Such unusual circumstances involve a combination of the following complicating factors that are relevant to all events.

- Multiple equipment failures and unavailabilities
- Deviations from operators' expectations, beyond their training and experience
- System conditions not addressed by procedures
- Time pressure
- Uncertainty and limited information
- Conflicting goals and shifting priorities

The above complicating factors frequently result in a significant mismatch between system behavior and operators' expectations; and it is this mismatch that frequently creates the need for operator actions. In retrospect, it may be quite obvious that the situational appraisals, goals, and action plans operators exhibited were inappropriate given the circumstances of the accident. Furthermore, it may be difficult to understand the tendency of operators to persist in implementing action plans given ineffective results. However, if we consider the seemingly illogical actions in the context of corresponding complicating factors, the actions often make perfect sense given the engineered and operational conditions.

The Three Mile Island (TMI) accident offers an example of how operators can be “made to fail” by combinations of operator mindset and unexpected, confusing conditions (see Kemeny et al. (1979) for a more detailed description of the accident). The accident occurred in Pennsylvania in March 1979 and has been the most serious commercial Nuclear Power Plant accident to occur in the United States, with significant reactor core damage and a small release of radioactivity. The accident started with a loss of heat sink (i.e., all secondary cooling), due to the loss of feed-water systems. The factors contributing to the unavailability of the systems included pre-existing misaligned valves and a maintenance tag that obstructed the position indicator of the valves. A pilot-operated relief valve (PORV) was then lifted automatically to relieve the pressure caused by the lost heat sink, releasing steam to a pressure relief tank (PRT). When the PORV failed to reclose as designed, the following factors acted together to complicate the situation and eventually lead to the accident.

- The actual position of the PORV was not displayed in the main control room. Although the valve was stuck open, after operators sent a close signal to the valve, they assumed that the valve was completely shut. As a result, the valve was open for more than two hours, causing water loss from the reactor vessel.

- As mentioned above, the steam released from the PORV was dumped into a PRT; however, the indicator for the tank water level, which could have shown an improperly functioning PORV, was on the back panel of the control room. The operators did not check the indication partly because they were overloaded. The workload and cognitive noise during the early and middle stages of the accident were excessive. For example, there were seven significant indications in the first 28 seconds following the unknown opening of the PORV.

- Another important factor was operators' understanding of the inexplicably rising water level in the pressurizer, as a consequence of the PORV being open. The rising level was an unanticipated, imminent phenomenon resulting from the loss of heat sink. The operators' understanding and responses to the phenomenon were conditioned by their naval submarine training, in which the dangers of a rising water level in the pressurizer were emphasized as it was indeed of more consequence and importance to operating a nuclear submarine than a nuclear power plant. This partly caused operators to fixate on the water level and overlook cues that could have indicated an ongoing Loss of Coolant Accident (LOCA). More importantly, driven by training emphasizing the need to respond to the rising water level in the pressurizer, the operators decided to switch off a safety system to avoid increasing the reactor pressure rapidly, which caused the reactor to be overheated and guaranteed there would be a meltdown of the reactor core.

The discussion above suggests that the cause of the accident was not simple and cannot be solely attributed to the operators. From a perspective of phenotype (i.e. what happened), the accident was a consequence of operators' inappropriate decisions and their overlooking cues that could have helped their diagnosis. From a perspective of genotype (i.e. why it happened), the operators' inappropriate decisions and actions were a result of their effort to create success, cope with complexity, and bridge the gap between their expectations and the evolving conditions by adapting their training and experience to the ongoing challenges. Their training and experience allowed them to perform skilled and speedy operations under normal circumstances; however, when applied in the wrong context, it led to inappropriate behavior with unsafe consequences. The operators did not knowingly commit an error; they were performing actions that seemed “correct” at the time. In general, this observation is consistent with those derived from other incident analyses and experiences described by training personal. That is, human error is not random and is conditioned by a conjunction of system conditions (e.g., equipment failures) and contextual factors, which in combination may be referred to as an error-forcing context or EFC (ATHEANA, 2000). By triggering internal psychological mechanisms, the EFC can lead to the refusal to believe evidence that runs counter to an initial misdiagnosis, or failure to recognize evidence, which can result in subsequent mistakes and ultimately, a catastrophic accident. This echoes the earlier discussion that attribution of failures to humans is an over-simplified explanation for failures. A proper understanding of human error requires a systematic and holistic point of view that accounts for the context. In fact, in recognition of the significance of context, it has been advocated by many researchers and practitioners to substitute the term “human error” with “unsafe actions” to avoid implications that the human was the cause of the problem or the human should take the blame for failures.

Error Classifications

A large number of taxonomies or classification schemes have been developed to describe error. For example, Meister (1971) divided human errors into four types based on where they originate: *operating errors*, *design errors*, *manufacturing errors*, and *installation and maintenance errors*. In probabilistic safety assessment (PSA) and human reliability analysis (HRA) for high consequence industries (e.g., the nuclear and aviation industries), human errors are often classified into three categories based on the relative timing of the errors and the accident sequence. The first category is *pre-initiator human errors*, which are faults that occur before the beginning of an accident sequence. The second category is *initiator human errors*, which are human actions contributing to the initiating event of an accident sequence. The third category is *postinitiator human errors*, which are faults that occur after an incident to aggravate the incident (IAEA, 1996).

The error taxonomies vary with respect to their theoretical orientations and practical concerns. Due to the complex nature of human error, it is difficult for a taxonomy to be comprehensive and reconcile specific contextual triggers with general error tendencies (Reason, 1990). That is, there is more than one way to classify an error; depending on the nature of the error and the tasks of interest. With a given application, a particular taxonomy may be preferred over others. The taxonomies range from simple binary classifications to complex hierarchical structures (Groth & Mosleh, 2012), and can be distinguished with respect to the following three levels at which the taxonomies approach the problem of classifying errors (Reason, 1990; Senders & Moray, 1991).

- Phenomenological or behavioral level. Taxonomies at this level classify errors in terms of the observable features of varied error phenomena. Although they can effectively describe *what* errors occurred at the superficial level of human behavior, these taxonomies do not provide insights with regard to *why* and *how* the errors occurred.
- Contextual level. Taxonomies at this level have an emphasis on the observable contextual triggering features that prompt an error; however, like phenomenological taxonomies, they provide limited insights as to why and how the errors occurred.

- Conceptual level. Taxonomies at this level are based upon conceptual considerations of the cognitive mechanisms involved in error production, and provide the greatest insight into the nature of the error.

Most of the available error taxonomies are useful only for retrospective error analysis, and few can be used to predict the imminent occurrence of errors (Senders & Moray, 1991). This is in contrast to the error classification schemes described in the following sections.

Errors of omission and errors of commission

One frequently used scheme classifies human errors into *errors of omission* (EOOs), which are instances in which actions that are necessary for a particular circumstance are not performed, and *errors of commission* (EOCs), which refers to errors in which incorrect actions are performed or an intended action is performed incorrectly (Swain & Guttman, 1983; Wickens, Gordon, & Liu, 1998). EOOs are often caused by distraction or diverted attention and are particularly prevalent in maintenance tasks. Generally, EOOs do not alter the trajectory of events and are correctable in that they merely result in time delays. In contrast, EOCs are often caused by inadequate training, poor instruction, or unrecognized hazards. EOCs do not always lead to undesirable consequences. Although the scheme was expanded by Swain and Guttman (1983) with two additional categories (*sequential errors*, which are actions performed out of the correct order, and *time errors*, which are actions performed too slow, too fast, or too late), only the first two categories have been widely used in HRA methods, such as the Technique for Human Error Rate Prediction (THERP) (Swain & Guttman, 1983).

Errors of omission, where an operator fails to perform some action, may be attributed to multiple causes. In an earlier chapter, there was discussion of mental lapses and the relationship between lapses and the brain's default network. A lapse refers to states in which an individual's conscious awareness is turned inward and as a result, they are less attentive and responsive to events occurring within the surrounding environment. Cheyne et al (2009) have linked errors of omission to mind wandering, attributing omissions to either momentary or prolonged lapses in attention. For instance, those exhibiting attentional deficits linked to Attention Deficit Hyperactivity Disorder (ADHD) show a greater incidence of errors of omission (Johnson et al, 2007). There is a natural tendency for task disengagement, with thoughts turning inward, and this disengagement tends to correspond to activation of the brain's default network.

Certain conditions make disengagement more likely, with some individuals being more prone to disengagement than others. Errors of omission can be expected to occur when disengagement results in either a failure to recognize cues prompting certain actions or there is a failure in task monitoring such that one loses track of what has and has not been done. Sustained task attention hinges upon the integrity of brain circuits associated with executive functions and contextual awareness, and conditions disrupting the integrity of these interrelated circuits increase the propensity for lapses and the associated errors of omission (Uddin et al, 2008). It has also been suggested that lapses may involve interference resulting from a failure to suppress task irrelevant, spontaneous activation of neural circuits (Helps, 2009). Thus, there is competition between task-oriented and task-irrelevant thoughts causing one to occasionally lose track of their ongoing activities. Taken together, it would appear that a primary source of errors of omission lies within the capacity of the brain to maintain sustained attention to a task, suppressing the tendency to drift into conscious states dominated by activity of the default mode network, with impediments to sustained attention (e.g. boredom, fatigue, distractions) increasing the likelihood of activities being omitted.

A second source of errors of omission involves situations where a routine or procedure has been learned, however there is a failure to adequately recall this knowledge from memory resulting in one or more steps being omitted. This situation will be most prevalent with routines and procedures for which there have been limited opportunities for practice with the learning being weakly engrained or there has been an extended duration during which there has been no practice and learning has begun to decay. It has been demonstrated that the ability to accurately recall a procedural routine depends on activity of a region of the brain known as the Supplementary Motor Area (SMA), during the period immediately following practice sessions (Tanaka et al, 2010). When currents were applied to the brain to suppress SMA activity immediately following practice, there was little benefit to practice, in contrast to a condition where currents were applied six hours subsequent to practice. It is suggested that the activity of the SMA serves to stabilize the memory representation within the primary motor area, and when this process is disrupted, the memory representation for a series of steps is more susceptible to decay. However, deficits in the ability to recall procedural routines have been linked to disruptions in the dopamine input to a region of the brain known as the caudate (Carbon et al, 2004). Dopamine serves to reinforce patterns of neural activation and in the absence of dopamine, memory representations linking series of steps should decay, leaving one more prone to omit steps or fail to complete a series of procedural steps. These findings suggest that the ability to accurately produce sequential patterns of behavior depends on the integrity of neural circuits that serve to establish and sustain the corresponding memory representations. When there is an inadequate opportunity to establish these routines or the routines are allowed to decay due to non-use, one might expect an increase in the incidence of errors of omission.

A third factor contributing to the occurrence of errors of omission involves failures of prospective memory. Prospective memory refers to the capacity to remember an activity to be performed sometime in the future. For example, talking to friends one evening, one may agree to call one of them the next morning. Prospective memory would involve the processes whereby one forms a memory representation for the activity that is to be performed the next day and then, recalls that memory the next day. When prospective memory fails, and one does not recall the intended activity, the resulting omission may be considered an error of omission. Brain imaging studies suggests that there are two somewhat distinct functional circuits associated with prospective memory (Burgess, Quayle & Firth, 2001). One serves to maintain an intention to perform the intended task, whereas the second provides the basis for realizing this intention. Consequently, activities that interfere with either function would be expected to produce prospective memory failures and associated errors of omission. For example, there are cognitive demands associated with sustaining an intention to carry out a future activity. Intervening demands due to other tasks or activities will diminish the capacity to sustain the intention. It is a common experience to realize that once a demanding situation has reached a conclusion, one or more routine tasks that would ordinarily have remembered have forgotten. Similarly, task demands at the time of the intended action can overshadow the intended activity. As a result, one may recall the intention, but being unable to carry out the intention, it is neglected and the corresponding actions omitted.

Slips (or lapses) and mistakes

A second widely used binary classification scheme distinguishes *slips* (or *lapses*) and *mistakes* (Norman, 1981; Reason, 1990). *Slips* (or *lapses*) are instances in which the intention is correct but a failure occurs when carrying out the associated activities (Reason & Mycielska, 1982). Simply put, slips (or lapses) are low-level errors of execution. Generally, a slip refers to an unintended deviation from a correct plan of action due to suboptimal attention allocation, whereas a lapse involves omitting part of a plan of action. As noted previously, both slips and lapses are unintended actions. Unlike the terms commission and omission, which are descriptive and reflect the impact of the error, the terms slip (or lapse) and mistake imply a causal mechanism.

Studies of slips and lapses date back more than a century when researchers, such as Meringer & Mayer (1895), Freud (1901), and several others, started to record, classify, and analyze slips of the tongue and pen to understand the underlying mechanisms for language production. Recent research concerning slips of the tongue shed light on the brain mechanisms underlying the occurrence of slips (Moller et al, 2007). In this research, subjects were exposed to an experimental procedure that is known to induce spoonerisms. A spoonerism is a slip of the tongue where the elements of a sentence are misplaced (e.g. instead of saying "go and take a shower," one might say, "go and shake a tower.") On trials in which subjects slipped and produced a spoonerism, there was heightened activation in the Supplementary Motor Area, a region of the brain associated with preparation of motor acts. This suggests that multiple motor acts were simultaneously prepared, with there being a failure to inhibit the inappropriate act, resulting in the spoonerism. Once the act, whether verbal or otherwise, has been initiated, there is often an almost immediate recognition of the slip, with a corresponding error-related response within the brain (Hiroaki et al, 2001).

In recent decades, extensive efforts have been devoted to expanding the studies to slips of action with a broader scope – understanding the organization of human performance and the role of consciousness in guiding action. In these studies, different classification schemes have been proposed for action slips. For example, Norman (1981) categorized slips into three groups (i.e., intention formation, activation, and triggering) based on the sources of the slips.

Slips may be distinguished from other forms of error with respect to the nature of the tasks and the mental and physical conditions that promote the occurrence of slips (Reason & Mycielska, 1982; Reason, 1990, P8).

- Slips occur during the execution of highly skilled or habitual tasks that require little continuous conscious monitoring. The likelihood of committing a slip increases with our proficiency at a particular task. This paradox can be explained by looking at the quantity and type of errors made by the expert and novice. Most of the errors made by the novice arise from a lack of competence and take unpredictable forms determined by the idiosyncrasy of the novice. As the novice becomes increasingly skilled at a task, increased automaticity in carrying out the details of the task diminishes the demand upon conscious attention. As a result, the novice begins to make fewer errors; however, their errors become more predictable in the sense that they tend to take the form of slips. That is, slips are the price we pay for the automaticity that allows us to perform routine tasks efficiently.
- Slips are most often triggered by the context. Routines are normally carried out in familiar environments where our vigilance is at a minimum. Environmental cues, such as

similarities in locations, actions, purposes, and expectations, can elicit well-organized segments of skilled actions that we habitually carry out in these circumstances. The actions are suitable for the context most of the time, but not when changed circumstances require some alteration of normal practice, or when new goals demand the modification of existing routines. Although the discussion here concerning context supports the notion, stated earlier, regarding the importance of context in understanding human error, it should be noted that contextual factors that are irrelevant to inducing automated behavior, such as noise, heat, pressure, and so on, do not seem to be important.

- Slips are often provoked by external distraction and/or internal preoccupation. Automaticity in performing routine tasks allows conscious attention to be directed to matters that are unrelated to the ongoing activity. Divided attention often occurs because the performance of routine tasks in familiar surroundings requires little vigilance. When inadequate attention is allocated to monitoring perception and action, inappropriate schemas or pre-established knowledge structures are likely to be activated by contextual factors and gain control resulting in biased perceptions or wrong impressions, which in turn lead to actions that are clearly recognizable as belonging to some other context or activity, but are inappropriate given the current intentions.

The effects of slips and lapses may be detected and corrected at various stages, ranging from the initiation of activity till the point when action departs from the plan. Many slips may be caught either by the perpetrator or sometimes with active cooperation of an observer. However, slips and lapses are often unrecognized, or in Reason's words (1990), remain "latent" for long periods of time, awaiting windows of accident opportunities (e.g., weaknesses in safety barriers) to reveal themselves. The misaligned valves and maintenance tag previously mentioned in the TMI scenario are good examples. When these chance events were combined with other operating circumstances, they had a significant impact resulting in a system breakdown.

Mistakes arise from the formation of an incorrect intention, which leads to an incorrect action sequence, although the actions may be consistent with the wrong intention. Compared to slips, mistakes are high-level errors with a substantial cognitive component and thus, mistakes are more resistant to detection and correction. Mistakes often occur at Rasmussen's (1983) level of knowledge-based processing (see discussion below) and result from failure to formulate a correct decision due to human processing limitations, incorrect knowledge, inadequate analysis, or planning failures.

The neural mechanisms that may be implicated in mistakes are much broader than those for slips. However, quite often, the origin of mistakes involves erroneous inferences concerning the current context, and while actions may be appropriate for the assumed context, they are not appropriate for the actual context. Brain imaging studies suggest that the recognition of context involves two distinct processes (Wan et al, 2011). One involves the perception of patterns within the environment and mapping of these patterns to a known context, with these operations linked to the precuneous region of the brain's parietal lobe. The second involves execution of actions based on recognition of a context, with this operation linked to the caudate nucleous of the basal ganglia. Failures associated with either of these operations provide the basis for making a mistake. Yet, mistakes may arise from other mechanisms as well. For example, memory for past events involves a process of construction in which memory traces are organized within the framework of current knowledge to construct a coherent recollection (Addis, Wong & Schacter, 2007). The same neural substrates that underlie the construction of memory for past events provide the basis for imaging and planning future events, with the hippocampus being a primary component and prospective memory involving additional engagement of frontal regions. Being a process of construction, memory is quite susceptible to erroneous formulations, with these errors in memory construction providing a basis for actions later characterized as mistakes. For example, given that the normal routine has a worker report to a specific location every morning, when trying to recall instructions from the previous day, it is easy to see how a memory might be formulated that erroneously infers that the worker should start the day at their usual location.

Skills, rules, and knowledge (SRK) taxonomy

Rasmussen's (1983) skills, rules, and knowledge (SRK) taxonomy provides another framework for human error classification based on the different types of information processing involved. According to the SRK taxonomy, operators' behavior can be classified into three categories based on the levels of cognitive control: *skill-based behavior* (SBB), *rule-based behavior* (RBB), and *knowledge-based behavior* (KBB). SBB consists of smooth, automated, and highly integrated patterns of action that are performed without conscious attention. It is usually based on feedforward, rather than feedback, control. A typical example of SBB is operators typing on a keyboard without visual support. RBB consists of stored rules derived from procedures, experience, instruction or previous problem solving activities. It is goal oriented, but does not require reasoning. Actions are directly triggered by familiar perceptual cues in the environment. KBB requires analytical reasoning and consists of deliberate and serial search based on an explicit representation of goals and a mental model of the functional properties of the environment. The SRK taxonomy–based scheme classifies human error into *skill-based*, *rule-based*, and *knowledge-based* errors. It should be noted that the slips/mistakes classification discussed earlier is closely related to the SRK classification. Slips are skill-based because they occur with well-practiced activities and are caused by misapplied competence. Mistakes, in contrast, are largely confined to the rule- and knowledge-based domains. In the rule-based mode, an incorrect diagnostic rule leads to incorrect intention. In the knowledge-based mode, incorrect intention results from excessive demands on operators' information-processing capabilities.

Summary

There is a compelling need to study human error. This is made clear by the fact that the functional failures of a human-technical system depend in many ways on the performance of its human operators, maintainers, and management. Particularly, some of the failures can, and sometimes do, result in devastating accidents of enormous size and cost. A careful analysis of human error as a psychological, behavioral and neurophysiological phenomenon can provide important clues about the underlying structure and organization of human performance and guide us to make useful theoretical inferences about the mechanics of various mental processes. Such clues and inferences will advance core theories and models underpinning the field of human behavior. They also offer important, practical utility in effectively reducing the adverse effect of error and the associated cost in operational settings.

The topic of human error is not simple and the progression in this field is slow. Despite extensive research, our knowledge of human error is still limited due to its complex nature. This is evident by the remarkable challenge in reaching a general consensus regarding the term "human error," not to mention the significant controversy in answering important questions such as how and why does error occur, and what factors contribute to the occurrence of error?

As discussed in the previous sections, human error is the product of the dynamic interaction between human, and systems or tasks to achieve a goal in a specified context; therefore, to understand error, it is important to examine all the elements involved in the interactive process. One important notion here is the strong dependence of human performance on specific aspects of the context. That is human error is due to the influence of the surrounding conditions on normal performance, rather than error mechanisms, and hence, is exogenous rather than endogenous. Analysis of error should take into consideration the relationships between error and particular environmental, contextual factors as well as the interactions among the contextual factors. With the benefit of hindsight, the right action or process may seem crystal clear. However, if we put seemingly illogical actions in the context of the surrounding, complicating factors, actions seem sensible given the engineered and operational setting in which they were planned and executed. The implications of the notion that human error is context-driven include: (1) we can reduce the likelihood of error occurrence by improving the conditions under which people work, and (2) error data collection should focus on good descriptions of the circumstances in which observed problems occur.

There has been efforts to formulate human error theories and models, and discover so-called human error mechanisms. This notion is misleading in that it implies that (1) there are two categories of human performance – *incorrect* and *correct* performance, and (2) each category of performance is governed by a separate, independent set of cognitive mechanisms. There is no evidence to support such an implication. Both incorrect and correct performance are produced by the same set of cognitive and neurophysiological mechanisms, which underlie all human performance. As such, the study of human error is actually the study of human performance. There is no need for theories or models specific to human error as an adequate model for correct human performance must be able to account for error; conversely, an adequate explanation of error must begin with some understanding of correct performance (Reason & Mycielska, 1982). In summary, we need a better theory or model of human performance as it relates to the social, organizational, and engineered context in which people live and work.

Acknowledgement

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contractDE-AC04-94AL85000.

References

Addis, D.R., Wong, A.T. & Schacter, D.L. (2007). Remembering the past and imaging the future : Common and distinct neural substrates during event construction and elaboration. *Neuropsychologia*, *45*(7), 1363-1377.

Anderson, J. R. (2000). *Cognitive Psychology and Its Implications* (5th ed.). New York, NY: Worth Publishers.

Bernstein, P.S., Scheffers, M.K. & Coles, M.G.H. (1995). “Where did I go wrong?” A psychophysiological analysis of error detection. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1312-1322.

Burgess, P.W., Quayle, A. & Firth, C.D. (2001). Brain regions involved in prospective memory as determined by positron emission tomography. *Neuropsychologia*, *39*(6), 545-555.

Busch, N.A., Dubois, J. & Van Rullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *Journal of Neuroscience*, *29*(24), 7869-7876.

Cacciabue, P. C., & Hollnagel, E. (1995). Simulation of cognition: Applications. In J.M. Hoc, P. C. Cacciabue & E. Hollnagel (Eds.), *Expertise and Technology: Cognition & Human–Computer Cooperation.* (pp. 55-73). Hillsdale, NJ England: Lawrence Erlbaum Associates.

Carbon, M., Ma, Y., Barnes, A., Dhawan, V., Chaly, T., Ghilardi, M.F. & Eidelberg, D. (2004). Caidate nucleous : Influence of dopaminergic input on sequence learning and brain activation in Parkinsonism. *NeuroImage*, *21*(4), 1497-1507.

Cheyne, J.A., Solman, G.J.F., Carriere, J.S.A. & Smilek, D. (2009). Anatomy of an error : A bidirectional state model of task engagement/disengagement and attention-related errors. *Cognition*, *111*, 98-113.

Churchland, M.M., Yu, B.M., Ryu, S.I., Santhanam, G. & Shenoy, K.V. (2006). Neural variability in premotor cortex provides a signature of motor prepration. *Journal of Neuroscience*, *26*(14), 3697-3712.

Dehaene, S., Posner, M.I. & Tucker, D.M. (1997). Localization of a neural system for error detection and compensation. *Psychological Science*, *5*, 303-305.

Fox, M.D., Snyder, A.Z., Zacks, J.M. & Raichle, M.E. (2005). Coherent spontaneous activity accounts for trial-to-trial variability in human evoked brain responses. *Nature Neuroscience*, *9*, 23-25.

Freud, S. (1901). *Psychopathology of Everyday Life*. London: Ernest Benn.

Gehring, W.J., Goss, B., Coles, M.G.H., Meyer, D.E. & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, *4*, 385-390.

Groth, K.M. & Mosleh, A. 2012. A data-informed PIF hierarchy for model-based Human

Reliability Analysis. *Reliability Engineering and System Safety*, *108*, 154-174.

Haynes, J.D., Sakai, K. Rees, G., Gilbert, S., Firth, C. & Passingham, R.E. (2007). Reading hidden intentions in the brain. *Current Biology*, *17*(4), 323-328.

Helps, S.K. (2009). Response variability in ADHD: Exploring the possible role of spontaneous brain activity. *Doctoral Dissertation*, University of Southampton.

Hesselmann, G., Kell, C.A., Eger, E. & Kleinschmidt, A., (2008). Spontaneous local variations in neural activity bias perceptual decisions. *Proceedings of the National Academy of Sciences*, *105*(31), 10984-10989.

Hiroaki, M., Hideaki, T., Noriyoshi, T. & Katuo, Y. (2001). Error-related brain potentials elicited by vocal errors. *NeuroReports*, *12*(9), 1851-1855.

Hollnagel, E. (1983). Position paper on human error. NATO Advanced Research Workshop on Human Error. Bellagio, Italy, 1983.

Hollnagel, E. (2007). Human error: trick or treat. In *Handbook of Applied Cognition* (2nd Edition). F. T. Durso (Ed), pp. 219-238.

Hollnagel, E. and Amalberti, R. (2001). *The emperor's new clothes, or whatever happened to 'human error'?* 4th International Workshop on Human Error, Safety and System Development, June 11-12, Linköping, Sweden.

Holroyd, C.B. & Coles, M.G.H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and its error-related negativity. *Psychological Reviews*, *109*(4), 679-709.

International Atomic Energy Agency (IAEA). (1996). Human reliability analysis in probabilistic safety assessment for nuclear power plants: A safety practice (IAEA Safety Series No. 50-P-10). Vienna, Austria: IAEA.

Ivergard, T., & Hunt, B. (2008). Models in process control. In *Handbook of Control Room Design and Ergonomics: A Perspective for the Future* (2nd edition), T. Ivergard & B. Hunt (Eds.), Boca Raton, FL: CRC Press, pp. 11–42.

Johnson, K.A., Robertson, I.H., Kelly, S.P., Silk, T.J., Barry, E. et al (2007). Dissociation of performance of children with ADHD and high-functioning autism on a task of sustained attention. *Neuropsychologia*, *45*(10), 2234-2245.

Keith, N. (2010). Managing errors during training. In *Human Fallibility: The Ambiguity of Errors for Work and Learning*, J. Bauer and C. Harteis (Eds) (Dordrecht: Springer, pp. 173-195.

Kemeny, J. G., and other 11 authors (1979). The need for change: The legacy of TMI – Report of the president's commission on the accident at Three Mile Island. New York: Pergamon Press.

Klein, G. A., Ross, K. G., Moon, B. M., Klein, D. E., Hoffman, R. R., & Hollnagel, E. (2003). Macrocognition. *Intelligent Systems, IEEE, 18*(3), 81-85.

Knoblich, G. & Sebanz, N. (2005). Agency in the face of error. *Trends in Cognitivce Sciences*, *9*(6), 259-261.

Letsky, M. (2007). *Macrocognition in Collaboration and Knowledge Ineroperability*. Paper presented at the 51st Annual Meeting of the Human Factors and Ergonomics Society, Baltimore, MD.

Lau, H.C., Rogers, R.D., Haggard, P. & Passingham, R.E. (2004). Attention to intention. *Science*, *303*(5661), 1208-1210.

Meister, D. (1971). *Human Factors: Theory and Practice*. New York: Wiley.

Nieuwenhuis, S., Ridderinkhof, K.R., Blom, J., Band, G.P.H. & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology*, *38*(5), 752-760.

Miltner, W.H.R., Braun, C.H. & Coles, M.G.H. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a generic neural system for error detection. *Journal of Cognitive Neuroscience*, *9*, 788-789.

Moller, J., Jansma, B.M., Rodriguez-Fornells, A. & Munte, T.F. (2007). What the brain does before the tongue slips. *Cerebral Cortex*, *17*(5), 1173-1178.

Norman, D. A. (1981). Categorization of action slips. *Psychological Review*, *88*(1), 1-15.

Norman, D. A. (1983). Position paper on human error. NATO Advanced Research Workshop on Human Error. Bellagio, Italy, 1983.

Nuclear Regulatory Commision (2000). *Technical Basis and Implementation Guidelines for A Technique for Human Event Analysis (ATHEANA)*. (NUREG-1624, Rev. 1). Washington, DC: US Nuclear Regulatory Commission.

Perrow, C. (1986). Complex organizations: A critical essay (3rd ed.). New York: Random House.

Rasmussen, J. (1983). Skills, rules, and knowledge: Signals, signs, and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC 13*, 257–266.

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

Reason, J. and Mycielska, K. (1982). Absent-minded?: The psychology of mental lapses and everyday errors. Englewood Cliffs, NJ: Prentice-Hall.

Scheffers, M.K. & Coles, M.G.H. (2000). Performance monitoring in a confusing world: Error-related brain activity, judgments of response accuracy, and types of errors. *Journal of*

Experimental Psychology Human Perception and Performance, *26*(1), 141-151.

Searle, J. R. (1980). The intentionality of intention and action. *Cognitive Science*, *4*, 47–70.

Senders, J. W. and Moray, N. P. (1991). Human error: Cause, prediction, and reduction, Hilsdale, NJ: Lawrence Erlbaum Associates.

Singer, W. (1993). Synchronization of cortical activity and its putative role in information processing and learning. *Annual Review of Physiology*, *55*, 349-374.

Soon, C.S., Brass, M., Heinz, H.J. & Haynes, J.D. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, *11*, 543-545.

Swain, A. D., & Guttman, H. E. (1983). Handbook of human reliability analysis with emphasis on nuclear power plant applications (NUREG/CR-1278, USNuclear Regulatory Commission).Washington, DC.

Tanaka, S., Honda, M., Hanakawa, T. & Cohen, L.G. (2010). Differential contribution of the supplementary motor area to stabilization of a procedural motor skill acquired through different practice schedules. *Cerebral Cortex*, *20*(9), 2114-2121.

Tassey. G. (2002). *The Economic Impacts of Inadequate Infrastructure for Software Testing*. National Institute of Standards and Technology, RTI Project Number 7007.011.

Uddin, L.Q., Kelly, A.M.C., Biswal, B.B., Margulies, D.S., Shedzad, Z. et al (2008). Network homogeneity reveals decreased integrity of default node network in ADHD. *Journal of Neuroscience Methods*, *169*(1), 249-254.

Ullen, F., Forsman, L., Blom, O., Karabanov, A. & Madison, G. (2008). Intelligence and variability in a simple timing task share neural substrates in the prefrontal white matter. *Journal of Neuroscience*, *28*(16), 4238-4243.

U.S. Nuclear Regulatory Commission. 2012. *NRC/EPRI Draft Report on Integrated Decision-*

Tree Human Event Analysis System (IDHEAS). Washington, D.C.: U.S. Nuclear Regulatory Commission.

Wan, X., Nakatani, H., Ueno, K., Asamizuya, T., Cheng, K. & Tanaka, K. (2011). The neural basis for intuitive best next-move generation in board game experts. *Sciecne*, *331*(6015), 341-346.

Weick, K.E. and K.M. Sutcliffe. 2001. Managing the Unexpected: Assuring High Performance in an Age of Complexity. Jossey-Bass: San Francisco.

Whaley, A. M., Xing, J., Boring, R. L., Hendrickson, S. M. L., Joe, J. C., LeBlanc, K. L., & Lois, E. 2012. *Building a Psychological Foundation for Human Reliability Analysis, NUREG-2114*. Washington, D.C.: U.S. Nuclear Regulatory Commission.

Wickens, C. D., Gordon, S. E., & Liu, Y. (1998). *An Introduction to Human Factors to Engineering*. New York: Addison-Wesley

Woods, D. D., Dekker, S., Cook, R., Johannesen, L., and Sarter, N. (2010). Behind Human Error (2nd Edition), Ashgate Pub Co.

Woods, D. D., Johannesen, L., Cook, R. I., & Sarter, N. B. (1994). *Behind Human Error: Cognitive Systems, Computers and Hindsight*. Crew Systems Ergonomic Information and Analysis Center, WPAFB, Dayton OH.

Table 1. Proximate causes for the failure of macrocognitive functions

Macrocognitive functions	Proximate causes
Failure of Detecting and Noticing	• Cues/information not perceived • Cues/information not attended to • Cues/information misperceived
Failure of Understanding and Sensemaking	• Incorrect data • Incorrect integration of data, frames, or data with a frame • Incorrect frame
Failure of Decision Making	• Incorrect goals or priorities set • Incorrect pattern matching • Incorrect mental simulation or evaluation of options
Failure of Action	• Failure to execute desired action • Execute desired action incorrectly
Failure of Team Coordination	• Failure of team communication • Error in leadership/supervision

Figure Captions

Figure 1. Human information processing model (Based on Card, Moran, & Newell, 1983; Ivergard & Hunt, 2008.