The Human as a Critical Component in an Adaptive Meaning Processing System

J.M. Flach
Department of Psychology
Wright State University
Dayton, OH 45431 USA
john.flach@wright.edu

Abstract - As automation has taken over more of the procedural aspects of work and as the complexity and the pace of change in the workplace has increased the human’s role with respect to safety has changed. The human’s ability to follow procedures without error is becoming less important, and the human’s ability to generate new procedures in response to changes or to situations that were not anticipated in the design of the system becomes increasingly more important. In other words, automation has taken over many of the routines that reflect inner control loops in advanced technical systems leaving humans to deal with the outer, adaptive control loops. The Information Processing Model, with its emphasis on internal processing limits, may not be the best framework for thinking about this new role. Instead, a “meaning processing” framework might better address the generative and creative demands of adapting to dynamic work environments.

Keywords: Human Error, Safety, Meaning Processing, Adaptive Control, Communication Channel, Control System

1 Introduction

Following a catastrophic accident, it is very easy to trace back along the events that led up to the accident to find actions that people took or failed to take that contributed to the adverse outcome. This fact, invites many people to incorrectly conclude that human error is a major ‘cause’ of these accidents. These accident analyses fail to see another important fact. They fail to see the many times when humans intervene to avert disaster. They fail to see the many times when humans deviate from the standard operating procedures and creatively intervene to insure safe and efficient operation of the system. This point is clearly illustrated with the example of “malicious procedural compliance,” that Vicente presents in the preface of his book on Cognitive Work Analysis [1]. In this example, a nuclear power plant team, who were frustrated by the fact that they received criticism from evaluators when they adaptively deviated from procedures during simulator evaluations, decided to follow a “work-to-rules” approach. The result was that the operators entered an infinite loop. As Vicente notes, “the evaluators were not amused.” They criticized the team for “malicious procedural compliance.” Unions and civil servants also use the “work-to-rules” strategy as a way to “strike” against management during labor negotiations [2].

As Rasmussen and Svedung note “work situations leave many degrees of freedom for choice by the actors, even when the objectives of work are fulfilled” [2]. These degrees of freedom are often resolved in the design process using assumptions framed around isolated tasks. These assumptions fail to fully appreciate the context of complex interactions among tasks that is the norm in many work situations. Similarly, these degrees of freedom can be difficult to appreciate in the hindsight following an accident. As Weick wrote, “people who know the outcome of a complex prior history of tangled, indeterminate events, remember that history as being much more determinant, leading ‘inevitably’ to the outcome they already knew” [3]. Reason also suggested how hindsight changes past indeterminacy and complexity into order, structure, and oversimplified causality [4].

This tendency to focus on human error and to overlook the creative ability of humans to manage complexity is reinforced by an information processing approach to human performance that is framed to identify internal information processing limits. For example, theories of perception focus on explaining illusions; theories of attention focuses on “bottlenecks” and “resource limits;” and theories of decision-making focus on explaining “biases.” Such an approach has lots to say about how the human cognitive system doesn’t work, but comparatively little to say about how it works. The ability of expert humans to cut through the complexity and to “zero-in” on the right action remains a deep mystery for basic cognitive science [5].

2 Communication channel versus control system

Two metaphors helped cognitive psychologists to escape from simple behaviorist models of cognition: 1) the communication channel and 2) the control system. Of these metaphors, the communication channel took a dominant role in the formation of the Information Processing paradigm. This metaphor matched well with the idea of a symbol processing system suggested by both linguistic and computer
metaphors. With all these metaphors, the focus is on open-loop, rule-based procedures for translating symbols to concepts. Meaning is considered to be a product of this translation process and the only basis for evaluating the "fitness" of this product is the adherence to normative rules in making the translation (e.g., deductive logic, Bayesian statistics, etc.). In this paradigm 'meaning' is a property of an internal representation with no grounding in either perception or action much less in the ensuing consequences. It is interesting that this lack of grounding tends to be seen as a "plus" by many cognitive scientists — who consider it essential to the "purity" of their science. Where "purity" typically translates to mean 'abstracted out from the messiness of natural systems.'

**TABLE 1**

<table>
<thead>
<tr>
<th>Approach</th>
<th>Information Processing</th>
<th>Meaning Processing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metaphor</td>
<td>Communication Channel Computer</td>
<td>Control System</td>
</tr>
<tr>
<td>Dynamic</td>
<td>Open-loop</td>
<td>Closed-loop</td>
</tr>
<tr>
<td>Meaning</td>
<td>Interpretation</td>
<td>Position in problem space (State)</td>
</tr>
<tr>
<td>Error</td>
<td>Deviation from normative process</td>
<td>Position relative to goal</td>
</tr>
<tr>
<td>Orientation</td>
<td>Internal Logic</td>
<td>External Demands Ecology</td>
</tr>
<tr>
<td>Level of Description</td>
<td>Algorithmic</td>
<td>Mechanistic</td>
</tr>
</tbody>
</table>

Would our view of cognition and the associated notions of "error" and "meaning" be different if the questions were framed using a control metaphor? With a control metaphor, perception and action tend to become significant aspects of the cognitive dynamic. A control system "actively" regulates a physical process. The actions of the control system are guided in a closed-loop fashion using information that is fed back and compared to a reference. The result of this comparison process is typically called the "error signal." In this context, "error" does not reflect an aberration or deviation from some normative process. Rather, error refers to a difference between two states (i.e., a goal or reference state and an estimate of the current state). Such differences are considered to be inevitable/natural features of any problem that a control system is designed to regulate. In this sense, "error" is often simply a position on a path toward a goal.

The term 'state' would be essential for addressing the issue of 'meaning' using a control metaphor. Essentially, the state is an index of position within a problem space. Where on the path toward the goal is the system? Choosing the indices that comprise a state description typically involves considering the dynamics of the problem being solved. This reflects not only the goals or references for success, but also the means for moving through the space (i.e., the constraints on perception and action). In other words, the states are the differences that make a difference with respect to the problem being solved — the states are the meaningful dimensions. In this context, 'meaning' is grounded in a functional analysis of the problem space. In a very real sense, a control system is organized around the states — it is organized around meaning. For this reason, we will refer to an approach to cognition based on a control metaphor as a "Meaning Processing" approach.

A final contrast between an Information Processing and a Meaning Processing approach is the orientation of the analysis process. An Information Processing approach tends to focus on the logic of internal processing mechanisms. This approach invites the scientist to look deeper into the neurophysiological processes to discover the logic of the internal mechanisms. A Meaning Processing approach tends to focus on the problem demands. In other words, a Meaning Processing approach tends to be 'ecological.' This orientation invites the scientist to look deeper into the problem to explore the possible solutions independent of any particular mechanism. In Marr's terms [6], the Information Processing approach is drawn toward algorithmic and mechanistic levels of description. The Meaning Processing perspective, on the other hand, is drawn more naturally to computational levels of description.

Table 1 summarizes some of the salient differences between the Information Processing and the Meaning Processing approaches to cognitive systems. The remainder of this paper will elaborate on the Meaning Processing approach with particular attention to the implications for managing safety in complex socio-technical systems.

3 The control metaphor

A simple servo-mechanism such as a thermostat is typically used as an icon for the control metaphor. However, this simple system is no less a 'mechanism' that the stages used in the communication metaphor. The point is to get past "mechanistic" images of cognition to address the creative, adaptive aspects of the phenomenon. To do this, it is essential to adopt the perspective of a control system designer or analyst. As Rasmussen et al. noted, "the issue is to consider the functional abstraction underlying control theory and to understand the implications of different control strategies on system behavior and design" [7].

3.1 Problem Formulation

Kirk's [8] introduction to Optimal Control Theory begins by citing the axiom, "a problem well put is a problem half solved." He then goes on to specify three requirements for the formulation of an optimal control problem:

1) A mathematical description (or model) of the process to be controlled.
2) A statement of the physical constraints.

It is important to keep in mind that Kirk is introducing optimal control as a computational tool for analysis, so his
focus is on "mathematical descriptions." For most complex, safety critical systems where the "human error problem" is an issue (e.g., aviation, nuclear power, medicine) we do not have reliable mathematical models of the problem to be controlled. Nevertheless, the spirit of these requirements reflects a control perspective. The first step is to understand the "process to be controlled." Tanabe once shared his frustrations in trying to communicate with psychologists interested in 'human error' the many factors contributing to an accident handling radioactive materials. He said, "It is very interesting for me to observe, concerning the accident, that many people talk about risk perception and human factors problems without investigating what kinds of risks were there in the working place and other aspects of the work" [9]. I wonder how many people are studying 'human error' with no deep understanding of or interest in the problems that the humans are wrestling with?

The development of models of the process to be controlled (whether quantitative or qualitative) begins with the choice of 'state variables.' As noted in the previous section - this is essentially a search for 'meaning.' It reflects an attempt to understand what differences make a difference. An excellent example of this search for meaning is the work of the Delft Man-Machine Systems group on the design of prostheses for children with missing arms. As Lunteren and Lunteren-Geritsen wrote: "this work started from the control engineer's view of a prosthesis as a technical system to be controlled, but finally evolved into a system consisting of wearer with his or her prosthesis in physical and social interaction with the environment" [10]. In the search for meaning, this research group conducted field studies following children through their daily activities. These field studies helped the researchers to appreciate cosmetic dimensions of the problem space, not apparent when the problem was framed in purely technical terms. Their view about what differences made a difference in this system evolved as a result of these field studies. While the insights from this search for meaning are difficult to formulate in a mathematical model -- there is not doubt that they had a major impact on the group's ability to design elegant solutions that were highly valued by the children.

In addition to the "states" (or meaning dimensions of the process), the second and third points in Kirk's formulation point to additional dimensions of meaning. Kirk's second point recognizes that there are boundaries within the "state space." These boundaries may reflect physical limits (e.g., maximum thrust for an aircraft), but they also might reflect information limits (e.g., limited range of different sensing systems).

Kirk's third point brings in the issue of value. In addition to the target or goal state, the costs and benefits associated with different paths through the problem space must also be considered in the design of a control system. There are typically many paths to the goal state, but some paths are more satisfactory (more nearly optimal) than other. Thus, the work analyst needs to compare and contrast different strategies for navigating the problem space in order to assess the costs and benefits associated with the alternatives.

In sum, a control theoretric approach means formulating the problem around issues of meaning; where meaning is framed as a property of a problem, NOT as a property of an individual's interpretation. The fundamental questions concern what differences are significant in terms of the functions, values, action constraints, and information constraints faced by the control system. In other words, the goal is to understand the problem in terms of 'means-ends relations' [7,11].

3.2 Adaptive control

Another weakness of the servomechanism as a representative of a control systems perspective is that it does not reflect the scale of the socio-technical systems where safety is a critical issue. These systems are very complex involving many degrees of freedom. So many, that it is practically impossible to consider all the dimensions that might make a difference in the operational context. Thus, no matter how committed we are to the search for meaning the best we can achieve is an asymptotic approach to understanding the controlled process. There will always be potentially critical factors not included in our models of the problem.

In addition to complexity, the fact that the world is constantly changing also sets a limit on the relevance of our models. The future that we are designing for is uncertain. Or in control theoretic terms, the processes we are trying to control are not stationary -- the problem dynamics change over time. This means that a solution that is satisfactory at one point may lead to instability at a later time.

The inability to consider all the variables or to unambiguously "see" into the future leads to the design of 'adaptive' control systems. That is, how can we design systems that can learn from their own mistakes? This generally involves designing systems with multiple nested feedback loops. Inner loops monitor motion through the state space relative to the goal, while outer loops are designed to evaluate and adjust the 'logic' of the internal loops. The logic might be evaluated against changing contexts (gain scheduling), performance ideals (model reference), or against new information obtained over time (self-tuning). Adjustments can include changing the goals, performance criterion, or the controller logic (e.g., changing the weights given to information about the state). Figure 1 illustrated some of the dimensions that might be reflected in the outer loops of an adaptive control system.

One lesson that emerges from almost any serious search for meaning is that context matters. In Suchman's terms, cognition is 'situated' [12]. This is nicely reflected in the prosthetics example cited earlier. The physical and social interactions with the environment were essential dimensions of the problem. In terms of Fig. 1 these dimensions would fit most naturally as part of the outer loops, while the technical
properties of the prosthesis would reflect the inner loop control problem.

It is important not to get too infatuated with the block diagrams (such as Fig.1) that we use to illustrate control systems [13]. Again, the metaphor is not with the 'mechanisms' of control, but with the orientation that control systems designers must take to problems. In designing and analyzing safety critical systems we must formulate the problems as adaptive control problems.

6 Implications for managing safety

In the context of an Information Processing approach, the safety problem is typically framed around the assumption that humans are a primary cause of major accidents. For example, Rouse & Rouse estimate that 'human error is the primary cause of 60 to 90 percent of major accidents in complex systems such as nuclear power, process control, and aviation [cited in Wickens, 14]. In this context, managing safety typically boils down to protecting the system against the 'limitations' that humans introduce into the system. This has caused many to look to 'automation' as a solution to the problem. However, experiences in aviation suggest that "automation does not eliminate human error, but changes its nature and possibly increases the severity of its consequences" [15].

It is tempting to draw comparisons with the U.S. Forestry Services' experiences controlling major fires. They have learned that managing forests requires more than simply stamping out forest fires. In fact, they have concluded that a successful program of stamping out forest fires can actually set the stage for major catastrophes. They have learned that fires are a natural part of the evolution of forests. Fires help to clear dead wood and brush and help to promote the germination of new life. If small fires are successfully prevented, the consequence is a build up of fuel within the forest. Eventually, as a result of a careless act or a lightning strike this fuel will ignite and will now burn at such great temperatures that germination is inhibited and the damage from the fire will be wide spread and long lasting.

The Meaning Processing approach offers a way to get past 'Smoky the Bear' approaches to safety that were designed to 'stamp out human error.' This approach assumes that humans are one of our best resources for designing adaptive systems. Most complex systems function efficiently because of the creative adaptations of the human components. Again, this is made apparent when "work-to-rules" policies are adopted. Humans are able to complete the design by adapting to dimensions of the control problem that were not anticipated by the people who designed the work standards and/or the technologies. When we attempt to manage safety by 'punishing' humans for deviating from standard procedures, we may be setting the stage for catastrophe, because we may be simultaneously cutting the outer loops necessary to support adaptation (preventing the germination of new, creative ideas).

A second implication of a Meaning Processing approach to safety is that we have to escape from the illusion of causal paths. As Powers observers "controlling means producing repeatable consequences by variable actions" [16]. The idea of tracing back along an activity sequence to discover the 'cause' of an accident has little or no value for understanding a control system. An alternative perspective on accident investigation is illustrated by Snook's [17] analysis of a friendly-fire incident and Woo and Vicente's [18] analysis of an e.coli
outbreak. Both analyses show the intricate web of constraints that give shape to an accident.

From the perspective of a Meaning Processing approach, managing safety is analogous to managing stability in a control system. The point is not to eliminate error, but to close-the-loop around error so that the system tends to converge toward desirable states. In control theoretic terms this typically comes down to a balance between gain and time delays [19]. But in more general terms this reduces to issues of sensitivity (gain) and timing. In other words, stability depends on getting the right information to the right place at the right time. In terms of Fig. 1 the goal is to make sure there is adequate information flowing in the various feedback loops.

This goal has been articulated by Rasmussen and Vicente as the concept of ‘ecological interface design’ (EID) [20, 21]. Consistent with Kirk’s first point, the EID approach starts by using the Abstraction Hierarchy to model the problem to be controlled (i.e., the work domain). The resulting model provides a guide for building representations that are meaningful. That is, building interfaces that accurately reflect the problem dynamics – that show the states in relation to the goals, the values, and the boundaries of safe operation.

5 Summary and conclusions

Any approach to safety that is based on tracing event trajectories to identify ‘causes’ that can be ‘stamped-out’ is doomed to failure. Not because it is difficult to find plausible ‘causes,’ but because it is too easy! Such approaches are motivated by a naive model of the dynamics of closed-loop systems. These approaches lead to simple, easy to understand, wrong answers. They may solve the blame problem, but not the safety problem.

In contrast to the search for ‘errors,’ we suggest that managing stability requires a search for ‘meaning.’ This requires that the safety problem be framed in the context of adaptive control systems. The goal should be cast in terms of managing stability of these systems, rather than in terms of eliminating errors. The first step is to identify the states of the control problem. That is, what are the dimensions that make a difference in terms of actions, information, and value? This entails examining alternative strategies in terms of their boundary conditions, information requirements, and their associated costs.

It is important that a control theoretic approach not be confused with the quantitative tools of control theory. The quantitative tools can be intelligently used to develop and evaluate insights about different control strategies. But it is paramount that the natural phenomena, NOT the tools define the issues of meaning. It is important that problems, like the need for a prosthetic limb, not be reduced to identifying parameters in an algorithm. The search for meaning must be grounded in the larger socio-technical context. The search must take an ecological, as opposed to technological, perspective. Quantitative tools should be used to support and test engineering judgment, they should not be used to replace engineering judgment.

The search for meaning is far from a simple answer to the safety problem. In fact, it is NOT an answer. It is simply a promising framework for posing the questions. The best we can hope for from such a search is an asymptotic approach to understanding the deep structure of complex socio-technical problems and an asymptotic approach to safety. The adaptive control theoretic context provides the best framework for carrying out the socio-technical debate about how to manage safety in a complex world. Safety is not a fixed target, but it is one dimension in an evolving complex space that we are simultaneously designing and exploring.

References


