Proceedings of the ASME 2012 11th Biennial Conference On Engineering Systems Design And Analysis ESDA2012 July 2-4, 2012, Nantes, France

ESDA2012-83015

DEEP STRUCTURE AND SMART MECHANISMS: DESIGNING PERSPICACIOUS SYSTEMS

John M. Flach
Wright State University
Dayton, OH, USA

ABSTRACT

A fundamental challenge in the design of any cognitive system is to support productive thinking and efficient control. Research shows that human problem solving can be greatly enhanced using representations that reflect the deep structure of problems. Further, research on human action shows that selectively constraining degrees of freedom can improve both speed and accuracy of performance. This talk will discuss how these two insights from the basic research literature can be incorporated into work analysis and interface design to enhance performance of cognitive systems. The goal is to design interfaces so that the deep structure of the problem is well mapped to the opportunities for action. A major challenge is to operationalize the basic constructs of *deep structure* and *smart mechanism* in terms of specific work domains. Examples from the medical and aviation domains will be used to illustrate how this challenge is being met.

INTRODUCTION

Over the last 30 years, there has been a gradual paradigm shift in how we frame problems of human system integration. The fields of Human Factors (HF), Ergonomics, and Human Computer Interaction (HCI) have been conventionally framed in terms of human limitations in information processing [1]. That is, the focus was on characterizing human limitations associated with perception, attention, memory, reasoning, and control, so that these limitations could be accommodated in the design of human-machine interfaces. In this context, these human limitations were often viewed as sources of human variability (i.e., human error) that were considered to be a major threat to system performance. The variability of the human was often seen as a serious 'weakness' relative to the reliability of modern information technologies. For example, Kantowitz and Sorkin [2] wrote:

> *Indeed, many human factors analysts believe that minimizing human error is the primary goal of any human factors design. If people never made errors, there would be little need for a science of human factors* (p. 30).

Research framed in the context of this conventional Information Processing (IP) paradigm typically was cast in the context of an open loop, dyadic semiotic system as illustrated in Figure 1A, where the human was treated as a symbol processor. That is, the interface was considered to be the stimulus/symbol/cause and the decision/action of the human was considered to be the response/interpretation/effect. Thus, the focus of research was typically to isolate open-loop transfer functions (e.g., determine the bandwidth) for the various stages of information processing (e.g., encoding, memory, decision making). In this context, even apparently dynamical aspects of human performance, such as Situation Awareness (SA) were typically cast as 'stages' in an information processing system [3].

In the context of the dyadic semiotic system the focus was on the relation between the surface structure of the interface (i.e., the symbol) and the interpretation of the human symbol processor. Research hypotheses in this paradigm were typically framed in terms of general surface properties of the interface (e.g., intergral versus separable displays), relative to human information processing limitations (e.g., parallel versus serial processing) and these hypotheses were often tested using generic tasks motivated by assumptions about the relevant information processes [1].

Limitations of this conventional IP/symbol processing paradigm became increasingly evident in the context of fault diagnosis in the nuclear power domain [4]. In this context, the ‘weakness’ of rule-based automation, in the face of unanticipated variability associated with a complex process became apparent, and it often fell to the human operators to discover and solve problems that could not be fully anticipated in the design of the systems (e.g., to diagnose faults related to failures of the automated control systems). Thus, human variability (now viewed as creative problem solving) was increasingly valued as a resource for coping with complexity. The implication for design was that there was an increasing concern to use automation in ways that supported human problem solving. Rather than protecting the system against human error, the emphasis shifted to how to more fully leverage the problem solving expertise of humans for coping with complex problems.

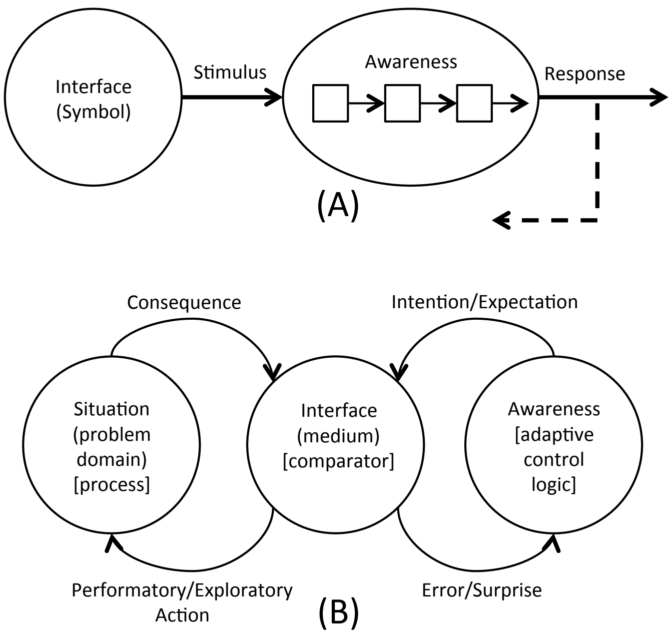

Figure 1. A comparison of a dyadic (A) and a triadic (B) image of the semiotic system for evaluating human system integration.

In order to understand creative human problem solving, attention shifted from research inspired by the information processing paradigm to other paradigms such as Gestalt [5] and Ecological [6,7,8] Psychology. These two paradigms focused more on the functional dynamics associated with ‘meaning processing,’ where meaning was derived from ‘grounding’ performance relative to the functional problem being solved. This grounding was reflected in constructs such as the ‘structural features’ of the problems discussed by Wertheimer and the ‘affordances’ discussed by Gibson.

For example, consider Wertheimer’s [5] view of thinking:

Thinking consists in envisaging, realizing structural features and structural requirements; proceeding in accordance with, and determined by, these requirements; thereby changing the situation in the direction of structural improvements, which involves:

that gaps, trouble-regions, disturbances, superficialities, etc. be viewed and dealt with structurally;
that inner structural relations – fitting or not fitting – be sought among such disturbances and the given situation as a whole and among its parts;
that there be operations of structural grouping and segregation of centering, etc.;
that operations be viewed and treated in their structural place, role, dynamic meaning, including realizing the changes which this involves;
realizing structural transposability, structural hierarchy, and separating structurally peripheral from fundamental features – a special case of grouping;
looking for structural rather than piecemeal truth (p. 235 – 236).

Brunswik and Gibson were among those to argue that the ‘structural relations’ and the ‘fitting’ that Wertheimer was concerned about reflected relations between the problem (i.e., the ecology or the situation) and the representation of that problem. They understood that in order to understand productive thinking it was necessary to consider the deep structure of the problem being solved, as well as the structure of the problem representation (whether internally as awareness or externally as an interface). They realized that the symbol processing had to be grounded in the functional practicalities of life – that it was not sufficient to understand the dynamics of awareness (e.g., information processing), it was necessary to consider the dynamics of situations (e.g., ecologies or work domains). This theme has been picked up by others and is reflected in labels such as situated cognition [9], embodied cognition [10]; and evolutionary psychology [11]. The common thread among these approaches is the need to include the situation constraints as a third component within the semiotic system.

Figure 1B illustrates a triadic semiotic system that includes the situation or problem dynamics as an additional intrinsic element in the meaning processing system. Although the triadic system includes two components that are labeled similarly to components in the dyadic system (i.e., Interface & Awareness), it is important to recognize that unique properties of these elements become salient in the context of the triadic system dynamic. That is, *the triadic system whole cannot be understood as the dyadic system plus the new situation component*. The triadic system challenges conventional assumptions about the role of interface and awareness, as well as assumptions about the global dynamics of situation awareness [12,13].

In the triadic semiotic system, the interface is no longer the ‘stimulus,’ but rather it becomes the medium through which functional constraints of the ecology can be specified. Gibson focused on the medium as optical flow. He hypothesized that structure within optical flow fields (i.e., optical invariants) specified important constraints in the ecology that were associated with the control of locomotion [14]. Brunswik [6], on the other hand, used his ‘lens model’ to describe the probabilistic relations between the medium (i.e., cues) and structures in both the ecology and the perceiver's beliefs about that ecology. Although they disagreed about the details of how to describe the constraints in the medium, Brunswik and Gibson both realized that understanding the constraints in the medium relative to structure in the problem ecology was critical to understanding the possibilities for skilled interaction with a functional environment. In other words, structure in the medium was critical to understanding the information coupling between situations and awareness.

The Awareness component also takes on a different character in the triadic semiotic system. The dyadic paradigm tended to treat the awareness component as a passive communication channel. Although feedback has been recognized as an important characteristic of cognitive systems [15], models of human information processing and the associated experimental paradigms have generally been framed using the logic of open loop causality [16]. The triadic paradigm, in contrast, suggests an adaptive control system metaphor that emphasizes the importance of the closed-loop coupling of perception and action in order to meet the functional demands of changing ecologies. For example, Gibson emphasized the essential role of action (e.g., looking) for information pick-up. In fact, optical flow does not exist without movement [17]. Gibson [18] introduced the constructs of performatory and exploratory action in recognition of the fact that action serves both control (e.g., goal satisfaction) and observation (e.g., information pick-up) functions.

The adaptive control metaphor demands that constructs of awareness be framed in relation to the functional demands of the control problem. As Conant and Ashby observed "there can no longer be question about *whether* the brain models its environment: it must " [19, p. 97]. The key point here is that the brain is designed through natural selection to leverage the opportunities in terms of information (e.g., optical invariants or cues) and in terms of functional action (e.g., affordances). It is not a logical, symbol processor, but rather it is a pragmatic system designed to solve the practical problems associated with surviving in an ecological niche (i.e., it is an adaptive regulator). In this regard, understanding the ecological niche, the dynamics of situations, provides a particularly important window into human awareness. In this context, situation awareness shifts from being internal components within an information processing element, to become an emergent property of the whole triadic semiotic system [12,13]. Situation awareness becomes the larger context for understanding both situations (e.g., affordances) and awareness (e.g., expertise).

The labeled arrows in Figure 1B reflect two aspects of the overall dynamic of the triadic semiotic: control and observation. Control is typically associated with performatory action and reflects the organization of behavior to satisfy or achieve some purpose or goal. This view was inspired by Rosenbluth, Wiener, and Bigelow's 'Cybernetic Hypothesis' [20] that suggested the servo-mechanism as a metaphor for human behavior. Observation on the other hand reflects the perceptual and sensemaking aspects of cognition. The terms in Figure 1B were chosen explicitly to reflect Pierce's [12] Abductive model of human reasoning. In this model, experience provides the source for hypotheses or beliefs about the world. The validity of these hypotheses is then tested pragmatically. That is, beliefs that lead to satisfactory outcomes are retained, while beliefs that lead to surprises or unsatisfactory outcomes are revised to reduce mismatches between the beliefs (i.e., internal model of the world) and the actual functional consequences.

It is important to note that the overall semiotic dynamic that results when control is coupled with and guided by an abductive logic is not consistent with the simple servomechanism that inspired early research in cognition [15]. As Rasmussen [21] has noted:

> *human activity, in a familiar environment will not be goal-controlled; rather, it will be oriented towards the goal and controlled by a set of rules which has proven successful previously…. The efficiency of humans in coping with complexity is largely due to the availability of a large repertoire of different mental representations of the environment from which rules to control behavior can be generated ad hoc. An analysis of the form of these mental models is important to the study of human interaction with complex man-made systems.*

In the adaptive control system described by Rasmussen a key to skilled performance will be matching the right models from the repertoire of mental representations (e.g., heuristics or control logic) to the right situations. The key will be to generalize appropriately from experience to situations. In other words, the capacity for control will depend on the match between structure in the heuristic or mental model guiding action (i.e., awareness) and the structure of the situation, as suggested by Todd and Gigerenzer's construct of ecological rationality [22]. In the triadic semiotic system, the interface (medium) becomes critical because the representation in the interface will have a critical influence on the generalizations that users will make. If the structure of the interface makes the functional structure of the situation or problem salient, then users are expected to make smart generalizations – to think productively. However, if the structure at the interface masks critical properties of the situation, then surprise and confusion are likely to be the result.

This leads to the fundamental premise guiding Ecological Interface Design (EID): that interfaces should be designed to reveal the functional structure (i.e., the deep structure) of the problem domain – so that operators will apply appropriate generalizations and heuristics for coping with the problem complexity [23]. It is not sufficient to simply consider human information processing. In addition, it becomes important to ground the information processes in the problem domain. Thus, the 'stimulus' is no longer the interface. Rather, the interface becomes the medium for representing the domain problem – which is the ultimate stimulus or ground for meaning processing. The actions/decisions/rationality of the human must be evaluated relative to the deep structure of the domain problem.

Note that human information processing limitations (constraints on awareness) remain a concern. However, these limitations are now framed relative to the deep structure of the work domain (constraints on situations). Thus, situation awareness is viewed not in terms of internal processing stages, but rather as an emergent property reflecting the fit between internal awareness (e.g., mental models or heuristics) and the functional situation dynamics. Also, note that in this context the validity of a heuristic is no longer gauged relative to context free logical prescriptions (e.g., deductive logic), but rather it is gauged relative to the pragmatic consequences (i.e., does it generally lead to satisfactory solutions).

Thus, the unique aspect of the EID approach relative to more classical user-centered design approaches is to treat the use-context (i.e., the work domain) as an intrinsic element of the semiotic system, rather than as an extrinsic environment surrounding a dyadic symbol processing system. In the following section, the construct of deep structure of a work domain will be explored and illustrated. Understanding what we mean by deep structure is essential to appreciating the EID approach. Following the discussion of deep structure, the implications for skilled control will be discussed in relation to the constructs of smart mechanism and coordinative structure. This will help to illustrate how representations of the deep structure of a work domain can be leveraged to reduce the computational demands of a problem.

DEEP STRUCTURE

In their book on human problem solving, Newell and Simon [24] write:

> *it would be perfectly possible for the psychologist to follow the route of the economist: to construct a theory of concept formation that depended on no characteristic of the subject other than his being motivated to perform well. It would be a theory of how perfectly rational man would behave in that task environment – hence, not a psychological theory but a theory of the structure of the task environment*. (p. 54)

In searching for the 'deep structure' of a problem (or in Wertheimer's terms the 'structural truths') we are searching for a 'theory of the structure of the task environment.' We are searching for an understanding of the rationality of the problem that is independent from any agent (either human or automation) that might need to solve it. For the puzzles and games studied by early researchers in AI, this deep understanding was typically achieved through 'state space' representations that showed all possible paths between an initial start state and a final goal state. This includes the states and the possible transitions from one state to another (i.e., operations or legal moves). The state space description provided a ground against which alternative solutions could be evaluated. For example, it allowed the identification of 'optimal' paths, in terms of minimizing the number of steps or minimizing costs associated with the various operations that needed to be applied in order to move from one state to another.

For simple games (e.g., tic-tac-toe) it is fairly easy to describe the state space. However, as problems become more complex (e.g., chess) it becomes increasingly difficult to fully describe the state space. Describing the state space becomes particularly difficult when we move from closed systems (e.g., games like tic-tac-toe or chess) to open systems (e.g., robots or complex work domains). An open system is one where the state depends, in part, on the environment or situation. For example, designing an autonomous robot for a controlled laboratory environment is much simpler than designing a robot that can be successful in a wide range of natural environments. As a system becomes more and more open to interactions with the larger environment, the complexity of the state space increases dramatically and it becomes increasingly difficult to fully enumerate the space. Thus, it becomes increasingly important to have a 'theory of the task' in order to differentiate the 'piecemeal truths' from the 'structural truths' of the problem space, in order to support productive thinking.

As Rasmussen [24] considered how best to support fault diagnosis in nuclear power plants, he realized that a 'theory of the task' was essential. Rasmussen realized that it was impossible to answer the question of what information was needed to diagnose a fault without a deep understanding of the nuclear power process. It became necessary to understand the 'states' of the plant and the 'operators' for changing those states. Rasmussen realized that understanding the dynamics of the nuclear power process was essential for appreciating the rationality of experts (e.g., why a particular strategy was preferred) and it was essential for making decisions about what information to display and how to organize it to best support those experts.

Rasmussen [25] introduced the Abstraction Hierarchy as a framework for building a theory of a functional task. The Abstraction Hierarchy includes five nested levels of means-ends constraints that reflected important general categories of task structure or constraint. I and others have written extensively about the Abstraction Hierarchy and how we think that it can help us to understand the deep structure of tasks [13,26,27,28,29]. However, in this article, I would like to avoid the jargon associated with the Abstraction Hierarchy and, instead, I would like to illustrate the concept of deep structure using concrete examples from our experiences in the medical and aviation domains.

Recently, we have been exploring the management of information in family medicine in relation to potential designs for electronic medical record systems (EMRs) [30,31,32,33,34]. In the course of this work, we have begun to focus on how the information from medical tests (e.g., blood analysis) can be used for decisions about treatment and care (e.g., relative to cardiac health). It soon became apparent that if we wanted to understand the rationality of medical decisions and if we wanted to provide appropriate information support, then we had to begin to learn about the nature of health. We had to begin to explore the relation between the data from the blood analysis (e.g., cholesterol levels) and the state of the patient's health (e.g., risk of cardiovascular disease – CVD risk). Thus, we began by asking domain experts about what variables do they attend to when making decisions about treating cardiovascular disease and we began to explore the medical literature for models of CVD risk.

Note that the goal of this search was not to identify the mental model of any particular physician. Rather, the goal was to find the best model of CVD risk. In other words, we were looking for a theory of the tasks of diagnosing and treating CVD. This search led us to regression models of CVD risk that were based on longitudinal studies of cardiac health [35,36,37]. The variables and weights in these models provided an empirical basis for relating the 'data' from blood analysis to the state of the patient with regard to CVD health (i.e., risk of cardiac event within next ten years), and it also provided an important link to the choice of treatment alternatives (e.g., whether to prescribe drug treatments).

Note that the domain experts were very helpful in guiding us to the right sources in the medical literature. However, although they were aware of the empirical models, they did not know the details of the models and could not articulate the logic of the models in a way to suggest innovative display solutions. At this point, we don't want to speculate about whether the domain experts' judgments about CVD risk would be in line with the predictions of the regression models. However, to the extent that they were consistent, we feel sure that much of this consistency would be based on tacit knowledge. That is, these judgments are based on a mental model that cannot be fully articulated. Thus, it would be difficult to discover this model with a work domain analysis based only on observations and interviews with practicing physicians.

The point for this paper is that the empirical regression models of CVD risk in the medical literature became critical to the theory of the task that has guided our choices about what information to include in a graphical display and about how to organize that information [33]. These regression models became our guide for finding 'meaning' in the data. They provided the theoretical basis for hypothesizing 'structural truths' that might support productive thinking. They have been important guides to our theory of the deep structure of the medical decision problem.

Another example comes from work in the aviation domain. In exploring alternative displays for landing an aircraft, we discovered that pilots used different strategies for landing on normal runways, than they used for landing on short fields (e.g., aircraft carriers). Although most of the pilots that we interviewed were aware of the different strategies, none could clearly articulate why one strategy (stick to speed) was best for normal runways and why the other strategy (throttle to speed) was best for short fields. The pilots had the procedural knowledge, but they did not have a deep understanding of the theory that motivated the different strategies.

As we explored further, we were led to more general questions about the functions of the throttle and stick relative to controlling speed and altitude. Note that either control can be used to regulate speed. In the stick to speed strategy, the pilot sets his throttle, and then flies using his stick to maintain a constant airspeed. In the throttle to speed strategy, the pilot sets his stick, and then flies using his throttle to maintain a constant airspeed.

Eventually, we learned that the functions of the controls could be more clearly differentiated in terms of energy constructs. We learned that the throttle determined the change in total energy available (i.e., the sum of potential and kinetic energy). When the throttle was increased, more total energy was available (i.e., the aircraft might either go faster, or higher, or both); when the throttle was decreased less total energy was available. We learned that the stick had essentially no impact on total energy, but that it determined the balance between potential (altitude) and kinetic (airspeed) energy. In other words, the stick position had an impact on whether energy was realized as changes in altitude or speed or both.

In this case, understanding the controls in terms of energy concepts provided a deeper understanding of the functions. This led to the development of a Total Energy Reference Path (TERP) display that when integrated with a tunnel in the sky display of glide path provides a graphical display that allows a much clearer differentiation between the functions of the throttle and the stick for controlling landing [38]. In the aviation domain, learning about the energy relations underlying the aerodynamics provided a deeper understanding of the task.

The construct of 'deep structure' refers to the need to develop a theory of a problem domain, situation, or ecology. This theory provides a basis for choosing what variables or what dimensions are important for relating the state of the problem to the functional goals. This theory provides a basis for differentiating whether different strategies will work and for evaluating the efficiency of various strategies. Thus, it also provides a ground against which the rationality of domain experts can better be appreciated, both in terms of the structural truths guiding experts and the limits or bounds of the heuristics guiding the experts.

Most importantly, however, a theory of the deep structure of a task can be a source of inspiration for display innovations that enhance expertise. That is, the theory of the deep structure of a problem domain can suggest ways to better organize information so that experts can 'see' structural truths that were invisible on conventional interfaces. A theory of the deep structure of a problem can guide the design of representations that shape the mental models of humans to better fit the functional demands of the problem. Thus, leading to more productive thinking.

SMART MECHANISMS

No matter how good our theory of the task or situation is the ultimate success of any control solution will ultimately be constrained by the information processing capacity of the control agent. For humans, we know that this capacity is severely limited. In fact, even for automated control systems the capacity is often quite limited relative to the complexity and dynamic demands of most work domains. It is important to appreciate that the issue is not simply the capacity of working memory, but that the information processing rate is also a function of the time it takes to access information. For example, in the case of sepsis with pre-mature infants, the time it takes to do the tests needed to diagnose the problem may exceed the critical time for effective intervention [39]. If the clinician waits for the lab results to confirm a diagnosis – the treatment will be too late to be effective.

For most complex situations, there is rarely enough time to collect and integrate all the information that might be relevant to a decision or control action. In dynamic problem domains, a controller that tries to take all the information into account before making a decision or taking an action will almost always be too late to be effective. Thus, in order to cope with complex problems, there is a requirement to parse or reduce the complexity in order to fit with both the internal (e.g., working memory) and external (e.g., limited window of opportunity) information processing limitations.

In fact, the strategies for landing described above are examples of how smart pilots reduce the complexity of the landing process to improve controllability. By fixing the throttle when flying a normal approach, the pilot reduces the complexity associated with interactions between the throttle and stick, so that attention can be focused on the stick-speed relation. Note that this only works if the throttle is set at the correct position (it is not arbitrary where the throttle is set). But if the throttle is set correctly, then the pilot can concentrate his attention on the stick control and will be able to achieve a very satisfactory solution to the problem of landing on a normal airfield.

In effect, the pilot behaves like one mechanism when landing on normal airfields (locking out the throttle and using the stick as the primary control), and he acts like a very different mechanism when landing on short fields (locking out the stick and using the throttle as the primary control). These are smart mechanisms to the extent that these strategies lead to satisfying solutions to the different situations. In each situation, the complexities associated with interactions of stick and throttle are reduced. However, the complexities are reduced in different ways to fit the different demands of the two situations.

The idea that a complex system might reduce complexity by locking out degrees of freedom to reduce the information processing demands was first suggested from research on skilled motor control [40]. Bernstein realized that controlling complex motor skills (e.g., throwing a baseball) could involve many degrees of freedom that potentially exceeded the information processing capacity of any centralized control system. He called this the degrees of freedom problem. However, he observed that skilled humans simplified the control demands by constraining different degrees of freedom to meet the demands of different situations (e.g., hitting a golf ball versus hitting a baseball). In other words, the skilled human solved the degrees of freedom problem by becoming different kinds of coordinative structures for different kinds of motor tasks. Thus, a coordinative structure is a heuristic solution fit to the demands of a specific task.

Inspired by Bernstein's intuitions, Runeson [41] contrasted the style of control in early robotics with that observed in biological systems. Whereas, early work in robotics tended to constrain movements around three orthogonal spatial axes (resulting in stereotypical 'robotic' motion and complex computations), biological systems tended to chose different axes for different tasks that better leveraged natural task constraints (resulting in smoother more natural motions and simpler computations). Runeson referred to solutions organized around the general dimensions most convenient for the engineers as 'rote mechanisms.' He referred to solutions organized around the special dimensions of specific task situations as 'smart mechanisms.' Thus, a smart mechanism reflects a solution that minimizes computational demands while satisfying the demands of the task.

In behaving as smart mechanisms, animals leverage natural constraints (e.g., the body's geometry, gravity, laws of motion) to reduce the control demands. For example, the trajectory of a limb can be constrained by locking a joint or by simultaneously adjusting multiple joints. The former strategy entails high computational demands as each joint and the interactions must be taken into account and coordinated precisely to yield a desired trajectory. The later strategy takes the locked joints out of the equation, greatly simplifying the computational demands.

The 'aiming off' strategy used by orienteers is another example of a smart mechanism. In orienteering, people using a compass and topographical map race from landmark to landmark based on directions obtained at each landmark. For example, the next landmark might be a bridge across a river that is a quarter mile through the woods northeast of the current location. Rather, than setting their course directly for the bridge, experienced orienteers will typically 'aim off.' That is, they set their course toward a point on the river that is closer than the bridge. On the one hand, this requires them to travel a bit farther than if they took the most direct path to the bridge, so it is sub-optimal relative to the shortest path. On the other hand, it is a computationally simple and robust strategy relative to the demands of following the compass without error to a precise point.

In navigating on a direct path to the bridge, there is a significant probability that due to errors along the route the orienteer will miss the bridge. Now when she sees the river she must guess which direction to follow the river (and how long to go before deciding to reverse directions). However, by aiming off, even if there have been errors along the route, the orienteer will know which way to follow the river to get to the bridge. In essence, with the aiming off strategy, the orienteer leverages the boundary constraint provided by the river to reduce the computational demands of locating the bridge. The position of the bridge is specified by the river, rather than by a series of compass computations. Thus, the aiming off strategy is an example of a smart mechanism.

In essence, a smart mechanism is a heuristic in the most positive sense of that word. It is a means to leverage local constraints of a problem to simplify the processing demands. It is a short cut. However, in leveraging local constraints, heuristics are necessarily bounded. They work only when the local constraints are present. In this case, rationality is bounded, but not due to internal processing constraints. It is bounded by the dynamics of the situation (i.e., it is situated) [42].

This is another reason that a deep theory of a problem is essential to designing effective representations. The deep theory helps to identify natural task constraints (e.g., energy balance relations) that can be leveraged to reduce the computational demands. A good representation is one that makes these constraints salient to the control agent in terms of salient patterns in the display. These patterns then effectively integrate information into meaningful chunks, helping to assure that the demands of the problem do not exceed the information processing capabilities of the human agents and helping the human to see both the value and the bounds of the local heuristics.

Note that the goal of simplifying or reducing the information processing demands is shared by both conventional approaches and the EID approach to display design. However, the key difference is that with the EID approach, the development of a theory of the task ecology becomes a guide to help to avoid the potential problem of trivializing a complex problem. The attention to the deep structure of the domain is motivated by the fear that if simplification is only guided by a focus on human information processing limitations, there can be a significant danger of reducing the complexity in a way that leads to brittle trivializations, rather than elegant solutions.

PERSPICACIOUS SYSTEMS

The adjective 'perspicacious' is defined as 'of acute mental vision or discernment.' This is an excellent description of the ultimate design goal for cognitive systems engineering. The goal is to design human-machine systems that have acute mental vision or discernment. The goal is to provide the decision support to doctors, pilots, and other people wresting with complex problems, so that they can be more perspicacious. To achieve this goal, we have to respect the limited capacity of human information processors. But this is not enough. In addition, we must appreciate the structural truths of the problem domains and situations that people are facing. In order to design effective representations for treating and diagnosing CVD, we must understand the deep structure of cardiovascular health. In order to design effective representations for piloting aircraft, we must understand the deep structure of aerodynamics. Thus, we must expand our image of the semiotic system to include the domain constraints as an intrinsic component of the system.

In designing interfaces for complex work domains, it is not sufficient to match the mental models of current experts. Rather, our challenge is to shape the mental models of experts so that they are more consistent with the strongest available theories and models of the problem domain. In medicine, this means to help doctors to better realize the promise of evidence-based practice, where their decisions are made in the context of the best empirical models from medical science. To do this, the data must be presented in a way that makes the relations to the medical models salient [33]. In aviation, this means to help pilots to go beyond procedural knowledge to better appreciate the physical aerodynamic principles of flight [37].

In this context, the interface is the medium for shaping alternative mental representations. Today's information and display technologies offer unparalleled opportunities to build alternative display representations. The challenge will be whether we can utilize the power of these technologies to make the critical variables salient and to organize them so that patterns that are produced illustrate constraints in the problem domain that can be leveraged against the complexity. The patterns have to correspond with deep structure in the domain and they have to be coherent within the perceptual capacities of the human agents.

Note that the goal here is slightly different than the goals for scientific visualization [43]. For scientific visualization the goal is to discover new patterns that challenge or extend current theories or knowledge. However, in designing for work domains, the goal it to build patterns that are guided by existing scientific theories. Thus, the goal is to bias doctors to make choices that reflect the best current medical theories – not to discover new theories. The goal is to bias pilots to make choices that reflect the best current aeronautical theories. Scientific visualization is about going from patterns in data to generate new theories. However, ecological interface design is about going from existing theories to generate new patterns to support practical decision-making and control. That is, to engineer the displays so that the patterns that result reflect the best existing theories of the domains. In the medical context, the goal is to help a physician to see the data for a specific patient relative to the best empirical models available from medical science. Thus, the patterns in the display should reflect the constraints in the medical models.

In this context, the focus shifts from the human as a source of error, to the human as a source for productive thinking. The human's capacity to recognize patterns becomes a major resource for dealing with problem complexity. The key is to make sure that the patterns perceived bias the human toward choosing the right heuristics for the situations; that they bias the human toward behaving as a smart mechanism.

In sum, the goal of cognitive systems engineering is to increase perspicacity. To do this requires both an appreciation of the constraints on human awareness and an appreciation of constraints within the problem ecology. The goal is to make sure that human discernment is shaped by the best existing theories of the deep structure in the problem ecology.

ACKNOWLEDGMENTS

These ideas have evolved through long term collaborations with Jens Rasmussen, Kim Vicente, and Kevin Bennett to explore general principles of cognitive systems engineering and ecological interface design. They have also benefited from specific collaborations with Matthijs Amelink, Max Mulder, and Rene van Paassen in the domain of aviation; and with Tim McEwen and Nancy Elder in the domain of family medicine. This work was partially supported through a subcontract to Aptima on a Phase 2 STTR from the Air Force Office of

Scientific Research titled: System for Information and Meta-information Portrayal of Lessons-learned (SIMPL) [P-9998-445].

Correspondences with regards to this paper should be addressed to John Flach, Department of Psychology, Wright State University, Dayton, OH. john.flach@wright.edu

REFERENCES

[1] Wickens, C.D. (1992). *Engineering psychology and human performance.* (2nd ed). New York: Harper Collins.

[2] Kantowitz, B.H. and Sorkin, R.D. (1983). *Human Factors: Understanding people-system relationships.* New York: John Wiley & Sons.

[3] Endsley, M. (1985). Toward a theory of situation awareness in dynamic systems, *Human Factors,* 37(1), 32-64.

[4] Rasmussen, J. and Lind, M. (1981). Coping with complexity (Risø-M-2293). Roskilde, Denmark: Risø National Laboratory.

[5] Wertheimer, M. (1959). *Productive thinking.* New York: Harper and Row.

[6] Brunswik, E. (1956). *Perception and the representative design of psychological experiments.* Berkeley: University of California Press.

[7] Tolman, E.C. & Brunswik, E. (1935). The organism and the causal texture of the environment. *Psychological Review*, 42, 43-77.

[8] Gibson, J.J. (1979/1986). *The ecological approach to visual perception.* Hillsdale, NJ: Erlbaum.

[9] Suchman, L.A. (1987). *Plans and situated actions.* New York: Cambridge University Press.

[10] Clark, A. (1997). *Being there. Putting brain, body and world together again.* Cambridge, MA: MIT Press.

[11] Barkow J.H., Cosmides, L., & Tooby, J.(1992).(eds.) *The Adapted Mind:Evolutionary Psychology and the generation of culture*. Oxford, Oxford University Press.

[12] Flach, J.M. (2009). The dynamics of experience. A search for what matters. Proceedings of the *European Conference on Cognitive Ergonomics (ECCE 2009),* p. 11 - 18, Helsinki, Finland (30 Sept - 2 Oct).

[13] Flach, J.M. & Rasmussen, J. (2000). Cognitive engineering: Designing for situation awareness. In N. Sarter & R. Amalberti (Eds.) *Cognitive Engineering in the Aviation Domain.* (p. 153 –179). Mahwah, NJ: Erlbaum.

[14] Gibson, J.J. (1958). Visually controlled locomotion and visual orientation in animals. *British Journal of Psychology*, 49, 182-194.

[15] Miller, G.A., Galanter, E., and Pribram, K. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart, & Winston.

[16] Jagacinski, R.J. and Flach, J.M. (2003). *Control theory for humans.* Mahwah, NJ: Erlbaum.

[17]. Gibson, J.J. (1966). *The senses considered as perceptual systems.* Boston, MA: Houghton Mifflin.

[18] Gibson, J.J. (1963). The useful dimensions of sensitivity. *American Psychologist*, 18, 1-15.

[19] Conant, R.C. and Ashby, R. (1970). Every good regulator of a system must be a model of that system. *International Journal of Systems Science,* 1, 89-97.

[20] Rosenbluth, A., Wiener, N. and Bigelow, J. (1943). Behaviour, purpose and teleology, *Phil. Sci.,* 10, 18-24.

[21] 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,* 13, 257-266.

[22] Todd, P. & Gigerenzer, G. (2003). Bounding rationality to the world, *Journal of Economic Psychology,* 24, 143 – 165.

[23] Rasmussen, J. and Vicente, K. (1989). Coping with human errors through system design: Implications for ecological interface design. *International Journal of Man-Machine Studies*, 31, 517-534.

[24] Newell, A. & Simon, H.A. (1972). *Human problem solving.* Englewood Cliffs, NJ: Prentice-Hall, Inc.

[25] Rasmussen (1986). *Information processing and human-machine interaction.* New York: North-Holland.

[26] Bennett, K.B. and Flach, J.M. (2011). *Display and interface design: Subtle science, exact art.* Boca Raton, FL: CRC Press.

[27] Flach, J.M., Mulder, M. & van Paassen, M.M. (2004). The concept of the situation in psychology. In S. Banbury and S. Tremblay (Eds.). *A cognitive approach to situation awareness: Theory and practice.* Aldershot, England: Ashgate.

[28] Amelink, M. (2011). *Ecological automation design: Extending work domain analysis.* Doctoral Dissertation, TU Delft, The Netherlands.

[29] Vicente, K.J. (1999). *Cognitive work analysis*. Mahwah, NJ: Erlbaum.

[30] McEwen, T. (2008). Creating Safety in the Diagnostic Testing Processes of Family Medical Practices. Unpublished Masters Thesis. Department of Psychology. Wright State University.

[31] McEwen,T.R., Elder,N.C. & Flach, J.M. (2011). Creating safety in primary care practice with electronic medical records requires the consideration of system dynamics. *Journal of Healthcare Engineering*, 2, 87 - 96.

[32] Elder, N.C., McEwen, T.R., Flach, J., Gallimore, J., and Pallerla, H. (2010). The management of test results in primary care: Does an electronic medical record make a difference? *Family Medicine,* 42(5), 327 - 333.

[33] Elder, N., McEwen, T.R., Flach, J.M. and Gallimore, J.J. (2009). The management of test results in family medicine offices: Complexity and quality. *The Annals of Family Medicine,* 7(4), 343 - 351.

[34] McEwen, T.R., Flach, J.M., and Elder, N.C. (2012). Ecological interface for assessing cardiac disease. *Proceedings of the ASME 2012 11th Biennial Conference on Engineering Systems Design and Analysis, ESDA2012*, July 2-4, Nantes, France.

[35] D'Agostino, Sr, R. B., Vasan, R. S., Pencina, M. J., Wolf, P. A., Cobain, M., Massaro, J. M., & Kannel, W. B. (2008). General cardiovascular risk profile for use in primary care: The Framingham heart study. *Circulation, 117,* 743-753. DOI: 10.1161/CIRCULATIONAHA.107.699579

[36] Ridker, P. M., Paynter, N. P., Rifai, N., Gaziano, J. M., & Cook, N. R. (2008). C-reactive protein and parental history improve global cardiovascular risk prediction: The Reynolds Risk score for men. *Circulation, 118,* 2243-2251.

[37] Ridker, P. M., Buring, J. E., Rifai, N., & Cook, N. R. (2007). Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: The Reynolds risk score. *Journal of the American Medical Association,* 297, 611-619.

[38] Amelink, H.J.M., Mulder, M., van Paassen, M.M., Flach, J.M. (2005). Theoretical foundations for total energy-based perspective flight-path displays for aircraft guidance. *International Journal of Aviation Psychology*, 15, 205 – 231.

[39] Crandall, B. and Getchell-Reiter, K. (1993). Critical decision method: a technique for eliciting concrete assessment indicators from the intuition of NICU nurses. *Advances in Nursing Sciences,* 16, 42-51.

[40] Bernstein, N.A. (1967). *The coordination and regulation of movements*. New York: Pergamon Press.

[41] Runeson, S. (1977). On the possibility of smart perceptual mechanisms. *Scandinavian Journal of Psychology*, 18, 172-179.

[42] Suchman, L. (1987). *Plans and situated actions: The problem of human-machine communication.* Cambridge: Cambridge University Press.

[43] Rosling, H. http://www.gapminder.org/