Aviation Human Factors Division
Institute of Aviation

University of Illinois
at Urbana-Champaign
1 Airport Road
Savoy, Illinois 61874

# A Systems Perspective on Situation Awareness I: Conceptual Framework, Modeling, and Quantitative Measurement

Alex Kirlik (University of Illinois)
and
Richard Strauss (FatWire Software)

Technical Report AHFD-03-12/NTSC-03-2

May 2003

Prepared for

Naval Training Systems Center
Orlando, FL

# A Systems Perspective on Situation Awareness I:

# Conceptual Framework, Modeling, and Quantitative Measurement

ALEX KIRLIK

Institute of Aviation, Departments of Psychology, Mechanical & Industrial

Engineering, and the Beckman Institute for Advanced Science & Technology

University of Illinois at Urbana-Champaign, Savoy, IL USA

RICHARD STRAUSS

FatWire Software, New York, NY USA

## ABSTRACT

Situation awareness (SA) has garnered much recent attention in the human factors community. SA inherently requires a systems perspective, as it concerns the degree of adaptive coupling between human cognition and an external environment. As such, we present an SA modeling approach giving equal attention to both the cognitive and external components of a human-environment system, in the realm of interface-mediated, uncertain judgment. The model allows SA in these contexts to be decomposed into seven measurable components. Importantly, we discuss how the model and measures map onto, and thus complement, theories of SA (e.g., Endsley), Human-Automation Interaction (e.g., Parasuraman, Sheridan & Wickens). and Naturalistic Decision Making (e.g., Klein). A companion article describes the first empirical evaluation of the utility of this modeling and measurement approach. Our central goal is to enhance theory and measurement of SA in support of design and training interventions.

# INTRODUCTION

Situation awareness (SA) has garnered much recent attention in the human factors and cognitive engineering communities (e.g., Adams, Tenney and Pew, 1995; Durso and Gronlund, 1999; Endsley and Garland, 2001; Wickens, 2002). This is not surprising, as information technology and automation in contemporary workplaces increasingly mediate the interaction between a human (or team) and the task environment which constitutes the target of work and performance. Although there are many dimensions to this phenomenon, and its role depends on the specifics of environment and task, few would disagree that continued scientific advances into understanding and supporting SA are sure to depend on advances in measurement. Salas, Prince, Baker, and Shrestha (1995) summed up the situation well: "a central problem in understanding situation awareness is the lack of well-developed measurement tools" (p. 131).

Our goal in this article is to present a technique for modeling and measuring a phenomenon we believe to lie at the crux of SA in a wide variety of task situations of interest to human factors: human judgment under uncertainty in conditions where judgment is mediated by a technological interface (e.g., a display of a remote system, situation, or environment). These assessments can be of either past, present, or future situations or events, and are based on the use of local, or proximal, information sources (such as those presented on an interface display) to infer a the existence of a situation, state or event present in a remote, or distal, task environment.

## *Theoretical Background and Purpose*

The methods we present have a systems, or ecological orientation, in that the intent is to capture *both* the cognitive *and* environmental (external to the performer) determinants of SA. While the majority of SA research has focused on the former, we agree with Pew (1995), who noted that "In order to adequately define SA we need to understand what we mean by a 'situation' and we need to know what it is about situations of which we must be aware" (p. 7). Consistent with this view, the methods we present are based on assumption that SA, as a relation between human cognition and an external situation, must be defined, modeled, and measured as such.

The measurement techniques we present, while somewhat new to human factors, and certainly new to the study of SA, are based on a long history of research and modeling in psychology, and more recent techniques borrowed from the weather forecasting literature. We present a model of the *judgmental aspects* of interface-mediated SA, resulting from the cumulative research of Brunswik (1956), who provided his lens model of judgment, the research of Hursch, Hammond, and Hursch (1964) and Tucker (1964), who provided a mathematical formulation for Brunswik's model, the research of Murphy (1988), who developed a diagnostic measure of weather forecasting skill, the research of Stewart (1990), who married the lens model equation with Murphy's skill measure, and finally, the research of Stewart and Lusk (1994), who supplemented the previous research by providing resources for describing the contribution of sensing and information processing technology for modeling interface-mediated judgment.

In this article, we present these systems-oriented, ecological techniques in the hopes they may provide an advance in the measurement and support of situation awareness. Importantly, we also provide information on how these methods connect to, and complement, related theories of

SA (e.g., Endsley, 1995a, 1995b), Naturalistic Decision Making, or NDM (e.g., Klein, 1999), and Human-Automation Interaction, or HAI (Parasuraman, Sheridan and Wickens, 2000).

## SA: COGNITIVE AND ECOLOGICAL APPROACHES

As the term itself suggests, situation awareness is a relational construct, requiring study of not only the contribution of cognition (e.g., perception, memory, knowledge, etc.), but also the contribution of the environment, as highlighted in Pew's comment in the previous section. Uncertain environments illustrate this point clearly, as the predictability of an environment places a constraint on the level of SA theoretically possible. If one's task was to maintain SA over a coin flipped in a locked room, outside one's presence, the upper bound (on average) for SA would be one-half. This can be learned prior to considering any psychological issues. In this case, the environment has irreducible uncertainty, and this uncertainty places a ceiling on SA.

### *Conceiving SA: An Awareness Focus*

Despite the observations above, it is fair to say that more attention has been paid to the awareness (cognitive) aspects of SA than to its environmental aspects. A proponent of research with this orientation has been Endsley, who has studied SA within the contexts of automation (Endsley, 1996), air traffic control (Endsley and Smolensky, 1998), and naturalistic decision making (Endsley, 1997), among others. Informally, she notes that SA concerns "knowing what is going on" (Endsley, 1995b, p. 36). More precisely, she defines SA as "the perception of the elements of the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" (Endsley, 1995b, p. 36).

The above definition describes SA as a three-level concept comprising elements of perception, comprehension, and projection. Endsley describes Level 1 SA as knowledge that results from the perception of disjointed "elements" in the environment, in which an element is equated with environmental objects and attributes. Level 2 SA, which is typified as a mental picture, is described as a comprehension of the situation, or an understanding of its current significance. Level 3 SA is described as the ability to predict future states of Level 1 elements.

Like Endsley, other researchers have been drawn to focus largely on the awareness, or cognitive side of SA, and thus have defined SA as a mental picture or an internal product. For example, in the context of land navigation, Wesler, Marshak, and Glumm (1998) equate SA to the contents of short-term memory. In aviation, Gibson and Garrett (1990) describe SA as a Gestalt-like appreciation of a situation, and Taylor (1990) describes SA as a "veridical model of reality" (p. 3-1),

### *Conceiving SA: An Ecological Focus*

Other researchers have promoted more systems-oriented, ecological definitions of SA. For example, Flach (1995) states that "SA defines the problem of human performance in terms of understanding the adaptive coupling between human and environment" (p. 153). He introduces the notion of *correspondence* to emphasize that "the human's awareness must correspond to the objective constraints of the situation" (p. 151). Smith and Hancock (1995) also promote an ecological focus, relying on Neisser's (1996) perceptual cycle to shape their definition. In their proposal, Smith and Hancock focus on elaborating the environmental components of SA. They do so by via Gibson's (1979) "invariant" concept, to denote the meaningful environmental features or constraints to which a person must adapt in order to achieve SA.

As discussed previously, Pew (1995) also points toward an ecological perspective on SA. Pew introduces two definitions, *ideal* and *obtainable awareness*. He defines *ideal awareness* as the awareness possible after all known information and knowledge requirements are satisfied. He then defines *obtainable awareness* as the level of *ideal awareness* possible after accounting for limitations in both knowledge and information. To return to our coin flipping example, imagine that a coin was flipped in one's presence, and one was asked for one's "awareness" of the outcome. In this case, both ideal and attainable awareness would be maximized, at least for those with the visual ability to view the outcome and the linguistic ability to report the outcome.

If the coin, however, were flipped in another, locked, room. and no relevant information pertaining to the outcome was available, both ideal and attainable awareness would fall to one-half. Imagine an example between these two extremes: If the coin was flipped in your presence but at some distance, ideal awareness would be perfect, and attainable awareness would depend on its distance and your visual discrimination ability. As you walk toward the coin in order to directly inspect it, obtainable awareness would rise to the level of ideal awareness.

## SA MODELING AND MEASUREMENT: A SYSTEMS APPROACH

Brunswik's (1956) ecological perspective on judgment under uncertainty is the initial basis for a model of the judgmental components of situation awareness (for a discussion of the relationship between Brunswikian theory and human factors, see Kirlik, 2000). Consider Figure 1, which depicts the basic components of both Brunswik's theory of judgment. and interface-mediated situation awareness.

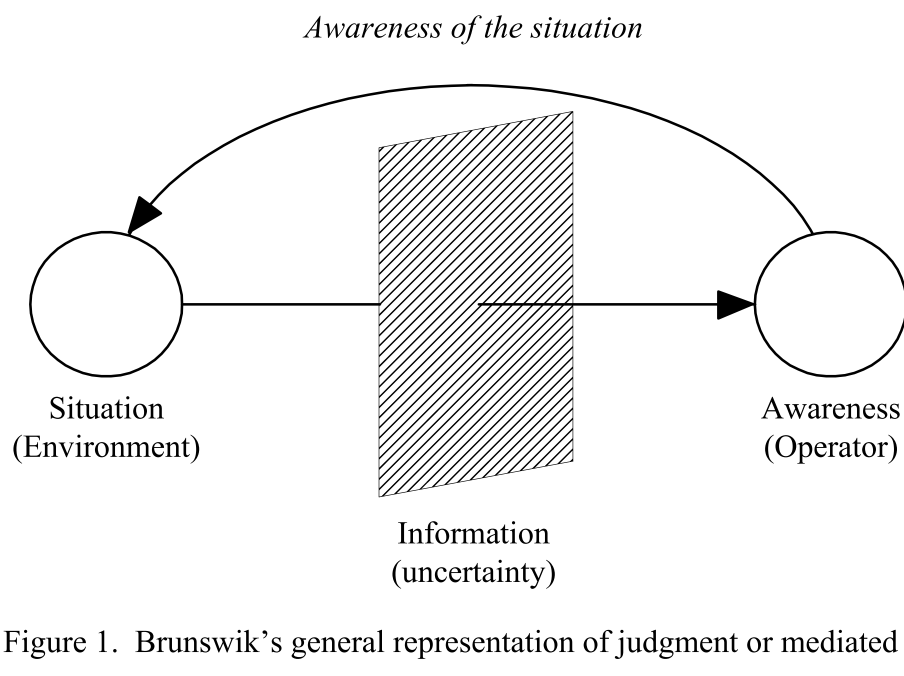


Figure 1. Brunswik's general representation of judgment or mediated SA

The left side of the Figure 1 depicts the situation, or what Brunswik referred to as the *environment*, and the right side of the figure depicts awareness, or what Brunswik referred to as the *organism* (we shall use "operator"). Mediating the situation-awareness relationship, or generally the environment-organism relationship, are both information and the uncertainty that accompanies it (i.e., the degree to which the information is capable of adequately specifying the state of the remote, or distal, situation or environment). Brunswik (1956) originally proposed measuring the quality of judgment in terms of a correspondence between the judged situation an the actual, environmental situation, measured by linear association, or bivariate correlation. Thus

the measure of achievement is obtained by correlating the operator's judgments with the true states of the situation being judged. Denoting this correlation coefficient $r_{YO}$, where the subscripts represent the contribution of the operator's judgment ($Y$) and the situation or "Object" of judgment ($O$), the higher the correlation ($0 \leq r_{YO} \leq 1$), the better the correspondence, the better the operator's achievement, and the better the SA.

*Limitations of the Correlation Coefficient*

Although the correlation coefficient provides a useful measure of correspondence, it has properties that limit its sensitivity, and thus utility, for measuring judgment quality (Cronbach and Gleser, 1953). Correlation captures only *shape* differences between two sets of variables, i.e., their shared pattern of ups and downs, without distinguishing the differences in either their *magnitude* or *scale*. A depiction of the correlation coefficient's insensitivity to both magnitude and scale is shown in the graphs in Figure 2. Note that the correlation is 1.0 in all four graphs.

Graphs (a) and (b) in show a difference in magnitude not captured by correlation. Graph (a) depicts a relationship between judgments (e.g., predicted distance of an approaching aircraft) and the true state of the situation (e.g., true distance of an approaching aircraft). In this case, average judgment is 11 (miles). Graph (b) shows another set of judgments for the same situation, but with an average judgment of 80 (miles). This inability of correlation to distinguish between these cases is evidence of the insensitivity of correlation to differences in magnitude. Graphs (c) and (d) in Figure 2 show a difference in scale not captured by linear correlation. In graph (c), judgments of a situation have a standard deviation of 5.9 (miles), while in graph (d), the standard deviation is 17.7 (miles), indicating the insensitivity of correlation to differences in scale.

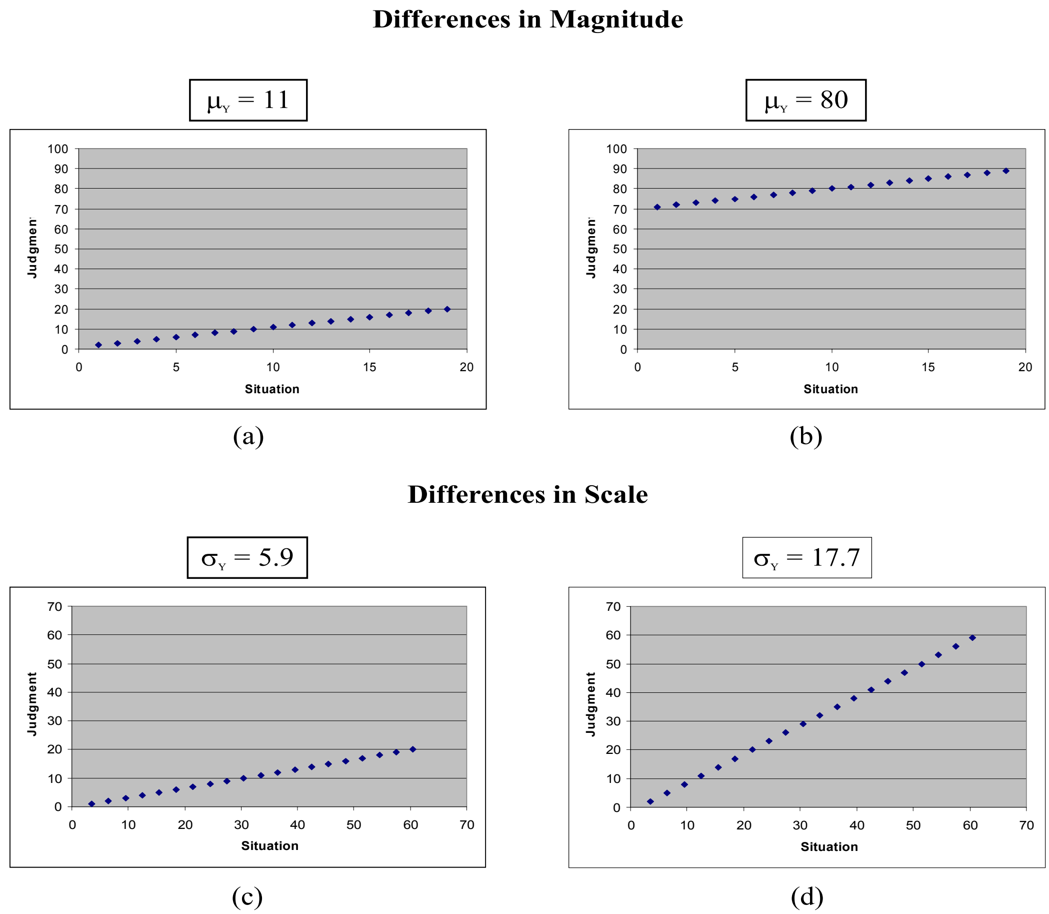


Figure 2. Insensitivity of the correlation coefficient to magnitude and scale

*Absolute Distance Measures of Judgment Quality*

These deficiencies of the correlation coefficient have motivated several researchers to look for more sensitive measurements of judgment correspondence. One alternative has been to look at the distance between data sets rather than their shared shape, a strategy often found in studies of meteorological forecasting (e.g., Murphy, 1988). Mean Square Error (MSE), a measure of the squared Euclidean distance between two data sets (Cooksey, 1996), has been regularly adopted for this purpose (e.g., Lee and Yates, 1992; Stewart and Lusk, 1994). MSE defines distance using the following equation,

$$MSE_Y = \left(1/n\right)\sum\left(Y_i - O_i\right)^2 \qquad \text{Equation 1}$$

Here, the two data sets are the judgments of the operator and the corresponding true states of the situation being judged. These two sets are used to form *n* pairs, where one element of the pair comes from each set ($Y_i$ and $O_i$ denote the i$^{th}$ judgment and i$^{th}$ true state respectively). When the judgments are perfect, MSE is equal to zero. As a replacement for the correlation coefficient, MSE would be unremarkable except that it can be partitioned into three distinct components representing shape, scale, and magnitude. Here, we present the decomposition proposed by Murphy (1988), introducing first, however, his skill score (SS) measure of judgment quality.

*The Skill Score as a Measurement of Judgment Quality*

To develop his decomposition of MSE, Murphy (1988) used the concept of *skill*, which he defined as judgment performance above chance. Chance performance is defined to be the degree of correspondence that would be obtained had a person always provided the same (constant) judgment based on the average, base-rate value of the situations being judged. In Equation 2, the quality of this standard is the $MSE_R$ that would be expected if the standard were always used.

$$MSE_R = \left(\tfrac{1}{n}\right)\sum\left(O_i - \overline{O}\right)^2 \qquad \text{Equation 2}$$

Here $\overline{O}$ is the mean, or base rate, of the observed event being judged. Deriving the skill score requires measuring the ratio between the MSE of the operator's judgment (Equation 1) and the MSE of the standard (Equation 2). This ratio is then subtracted from unity to create the skill score (SS). This relationship is shown in Equation 3. In this basic form, the skill score provides overall evaluation of the quality of an operator's judgments as compared to chance. When SS is

$$SS = 1 - \left[\tfrac{MSE_Y}{MSE_R}\right] \qquad \text{Equation 3}$$

positive, the operator's judgments are better than chance ($MSE_Y < MSE_R$); when it is zero, the judgments of the operator are equal to chance performance ($MSE_R = MSE_Y$); and when SS it is negative, the operator's judgments are worse than chance ($MSE_Y > MSE_R$).

*Murphy's (1988) Decomposition of the Skill Score*

Murphy (1988) developed the SS to enable the MSE to be decomposed. By substituting the equations for $MSE_Y$ (Equation 1) and $MSE_R$ (Equation 2) into the form of the skill score (Equation 3), Murphy (1988) showed how to derive the desired decomposition. A conceptual representation of his decomposition is presented in Equation 4:

Judgment Quality (SS) = [Shape (correlation) – Scale Error – Magnitude Error] (Equation 4)

Here, the Skill Score (SS) is partitioned into three components, and thus shape (correlation) is separated from errors associated with differences in magnitude and scale. The result is a more sensitive and diagnostic measurement than correlation alone (which measures only shape similarity). The scale error component has been called *Regression Bias*, as it measures whether the operator has appropriately scaled judgmental variability to situational variability. It is zero when the slope of the regression line predicting the observed events from the operator's judgments is 1.0 (Stewart and Lusk, 1994). For example, a submarine sonar technician with a regression bias might on average judge the range of an approaching enemy to be between 20 and 100 NM, when the actual interval of ranges is between 50 and 80 NM. A regression bias is a tendency to produce judgments in either a smaller or larger range than in the actual situation.

Finally, the magnitude error component of Murphy's Equation 4 has been called *Base Rate Bias* (Stewart, 1990). It measures the overall (unconditional) bias in the operator's judgments, thus diagnosing a tendency to over- or underestimate the judged situation. This bias equals zero only when the mean of the operator's judgments equals the mean of the judged states (i.e., the objective base rate, and is non-zero when the mean operator's judgment is too high or low).

Murphy's decomposition of the skill score provides a sensitive measure of the judgmental components of situation awareness, as it disentangles the joint contributions of shape, scale, and magnitude in the measurement of judgment quality. The mathematical decomposition is presented below in Equation 5, and its components are summarized in Table 1.

$$SS = (r_{YO})^2 - \left[ r_{YO} - \left( s_Y / s_O \right) \right]^2 - \left[ \left( \overline{Y} - \overline{O} \right) / s_O \right]^2 \qquad \text{(Equation 5)}$$

| Component | Name | Description |
|---|---|---|
| SS | Skill Score | A relative measure of "actual" judgment quality. |
| $r_{YO}$ | Correlation Coefficient | **Shape**—degree of linear association between judgments and situation. "Potential" skill in judgment. |
| $\left[ r_{YO} - \left( s_Y / s_O \right) \right]^2$ | Conditional/ Regression Bias | **Scale**—degree that standard deviation of judgments accounts for imperfect correlation; for the bias to vanish, $s_Y$ must be adjusted to equal $r_{YO}(s_O)$. |
| $\left[ \left( \overline{Y} - \overline{O} \right) / s_O \right]^2$ | Unconditional/ Base Rate Bias | **Magnitude**—degree that average judgment equals the base rate of occurrence in the situation. |

Table 1. Components of Murphy's (1988) decomposition of the skill score

*Augmenting the Skill Score with the Lens Model Equation*

The decomposition of the skill score, as shown in Equation 5 and Table 1, can improve the diagnosticity of the judgmental components of SA. Additional diagnosticity can be gained by taking this decomposition one step further by decomposing the correlation coefficient ($r_{YO}$). To do so, consider a more detailed depiction of Brunswik's lens model as shown in Figure 3.

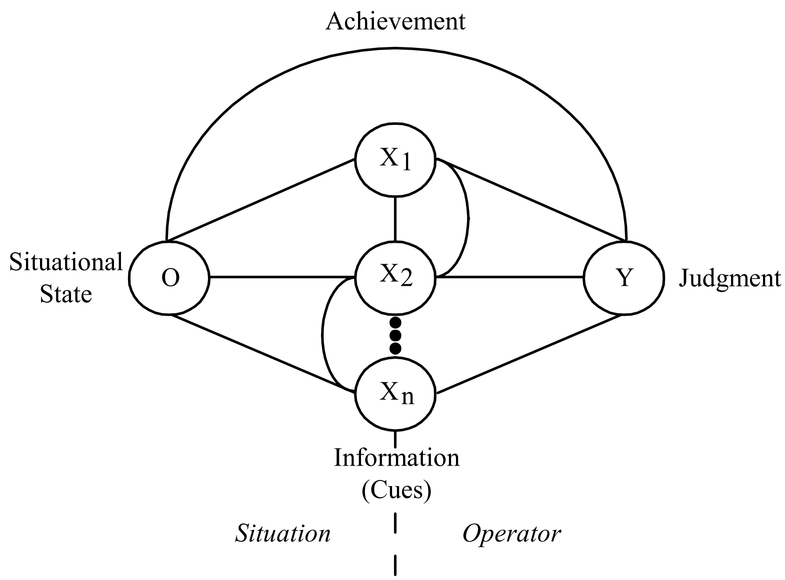


Figure 3. Brunswik's lens model.

The lens model shares the same configuration as Figure 1, yet depicts information as an enumerable set of cues, or items of information (labeled $X_i$'s). Recall that Brunswik measured achievement (the top arc in Figure 3) as a linear correlation, in our case, $r_{YO}$. Each relationship between a cue and a situation can be assigned an *ecological validity*, or the degree to which a cue informative about the situation. In addition, each relationship between a cue the degree to which it is relied on by an operator can be assigned a *cue utilization*. The lens model can be used to further decompose Murphy's Skill Score by examining the relations among these variables.

*The Lens Model Equation*

One of the most important extensions to Brunswik's lens model was the development of the lens model equation (Hursch, Hammond, and Hursch, 1964; Tucker, 1964). The lens model equation (LME) provides a mathematical representation of the lens model and partitions the overall correlation represented by the level of achievement or $r_{YO}$ into correlations related to ecological validities of cues, cue utilizations, the predictability of the environment, and the consistency with which an operator implements his or her judgment (cue-weighting) strategy.

At the basis of the LME are two parallel models, which represent the *Situation* side and the *Operator* side of the lens model shown in Figure 3. Both models are typically implemented with multiple linear regression models, but this need not be the case (e.g., see Rothrock and Kirlik, in press, for a discussion of this issue and an alternative formulation in terms of rule-based modeling). The situation model describes the overall correspondence between the cues ($X_i$'s) and the situation ($O$), and the operator model describes the overall correspondence between the cues ($X_i$'s) and the operator's judgment ($Y$). Based on these two models, the resulting decomposition of achievement is depicted conceptually in Equation 6.

$$(r_{YO}) = \text{Environmental Predictability x Consistency x Knowledge + Error} \quad \text{(Equation 6)}$$

*Environmental Predictability* (or equivalently situational predictability) is the correspondence between the cues and the situation. *Consistency* is the correspondence between the cues and the operator's judgments as reflected in the operator model. Thus, a lower correlation between the cues and judgment behavior is less than fully predictable based on knowledge of the cues.

*Knowledge* is the degree of correspondence between the outputs of the situation and the operator models. Outputs from these models represent the predictable aspects of the situation and operator based on their respective degrees of correspondence with the cues. This component is called *Knowledge* because it indicates the degree to which the operator correctly models the regularities of the situation, or weights the cues adaptively. *Model Error* in Equation 6 is the degree of correspondence between the unpredictable portions of the situation and operator models. Typically, this value is found to be marginal in human judgment (Brehmer, 1994).

$$r_{YO} = R_{O.X}GR_{Y.X} + C\sqrt{1 - R_{Y.X}^{2}}\sqrt{1 - R_{O.X}^{2}} \qquad \text{Equation 7}$$

As discussed previously, the decomposition shown in Equation 7 is accomplished with multiple linear regression. Thus, *Environmental Predictability*, *Consistency*, *Knowledge*, and *Model Error* are measured using multiple correlation statistics. The mathematical form of the LME is shown above in Equation 7, and its components are summarized below in Table 2.

| **Component** | **Name** | **Description** |
|---|---|---|
| $r_{YO}$ | Achievement | Correlation between judgments and situation |
| $R_{O.X}$ | Environmental Predictability | Correlation between situation and cues |
| G | Knowledge | Correlation between the model of the situation and model of the operator |
| $R_{Y.X}$ | Consistency | Correlation between cues and judgments |
| C | Model Error | Correlation between the residuals (or errors) in both models |

Table 2. Components of the lens model equation

## THE EXPANDED LENS MODEL

To integrate the benefits of both the LME's decomposition of achievement and Murphy's (1988) decomposition of the skill score (SS), Stewart (1990) presented a single equation showing how these decompositions could be combined. In Stewart's model, $r_{YO}$ from the SS in Murphy's decomposition (see Equation 7) is expanded using a partial form of the LME (one excluding the typically negligible C or error component). A conceptual form of Stewart's decomposition is presented below in Equation 8 and its mathematical form is presented in Equation 9.

The first three terms in Equation 8 are from the LME decomposition of $r_{YO}$. These three terms are respectively translated into the first three measures shown in Equation 9. The last two terms in Equation 8 (*Scale Error* and *Magnitude Error*) remain unchanged from Murphy's original decomposition of the SS. They are translated into the last two measures of Equation 9:

SS = Env. Predictability x Knowledge x Consistency – Regression Bias – Base Rate Bias (Eq. 8)

$$SS = \left(R_{O,X} G R_{Y.X}\right)^2 - \left[r_{YO} - \left(\frac{s_Y}{s_O}\right)\right]^2 - \left[\frac{\left(\bar{Y} - \bar{O}\right)}{s_O}\right]^2 \qquad \text{Equation 9}$$

*The Expanded Lens Model*

When considering the structure of the model presented in Equations 8 and 9, Stewart and Lusk (1994) recognized that *Environmental Predictability* (the correlation between the cues and a situation) could be further decomposed into two aspects: (1) the process that extracts data from the situation and (2) the process that transforms extracted data into cues available to the operator. These two processes represent a sequence typical in many systems with interface technology.

For example, in a submarine the first process is represented by the sensing technology that gathers raw data from the underwater environment. The second process is represented by the fusing and display technology that transforms this data into the cues presented to the operator. Note that in this example, noise in the sensors or underwater environment can be passed on to the raw data, which can ultimately be passed on to the cues.

Decomposing *Environmental Predictability* into a two-stage sequence provides a more diagnostic description of the situation-information relationship. Moreover, it can help to identify constraints, such as malfunctioning sensors, that attenuate higher degrees of correspondence. This sequence of “technological information processing” is presented above in Figure 4.

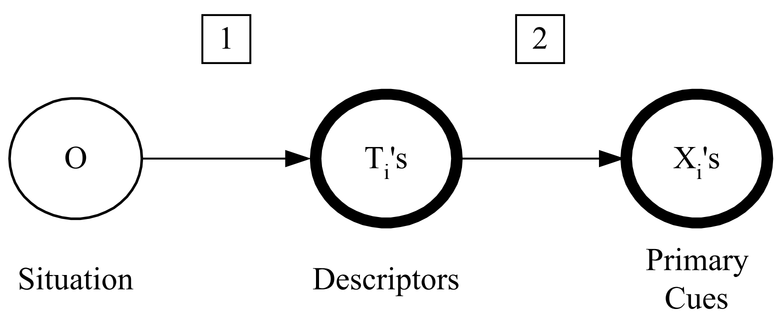


Figure 4. Two-stage sequence of technological information processing.

As shown in Figure 4, the two stages are labeled [1] and [2]. Stage [1] depicts the extraction of data or *descriptors* (labeled $T_i$’s) from the situation. In Stage [2], these descriptors are transformed into the primary cues provided to the operator via interface displays. The thickened outlines around both the descriptors and the primary cues indicate their status as sets. Note that the situation and primary cues ($O$ and $X_i$’s, respectively) correspond to the situation and cues depicted in the traditional lens model.

In a similar manner, Stewart and Lusk also recognized that *Consistency*, a measurement of the correspondence between information (cues) and operator judgments, could be further decomposed into two processes: (1) the process that an operator uses to acquire information, and

(2) the process that the operator uses to transform that information into a judgment. Akin to the previous sequence of technological information processing, errors or noise in the first process can be passed on to the second and ultimately constrain judgment quality.

Decomposing *Consistency* into a two-stage sequence increases the diagnosticity of the description of the information-operator relationship. Moreover, it can help to identify human information processing constraints that attenuate higher degrees of correspondence. The two-stage sequence of "operator information processing" (IP) is presented above in Figure 5. As shown in Figure 5, the two stages of operator IP are labeled [1] and [2]. Stage [1] depicts the operator's acquisition of *secondary cues* (labeled $U_i$'s) from primary cues. In Stage [2], these secondary cues are transformed into a judgment. The primary cues in Figures 4 and 5 are the same.

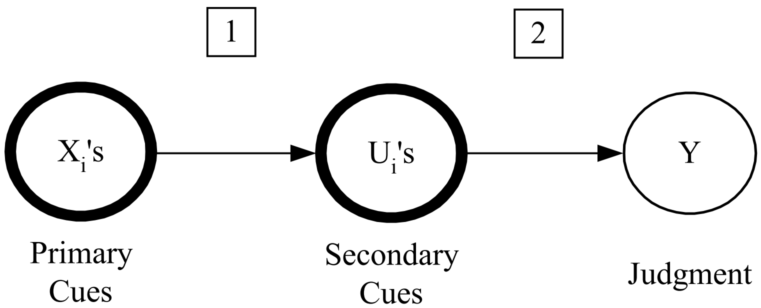


Figure 5. Two-stage sequence of operator information processing.

Operator IP shown in Figure 5 represents a sequence common to many IP models of SA. For example, Endsley's (1995b) three-level description characterizes SA as (1) perception of elements in the situation, (2) comprehension of the situation based on a transformation of those elements, and (3) a judgment of the future states of those elements. When judgments are made of future states (e.g., the likelihood of a collision with an approaching aircraft), the two-stage description in Figure 5 maps to Endsley's first and third levels, where secondary cues ($U_i$'s) are mapped to Endsley's Level 1 SA, and the operator's judgment ($Y$) is mapped to Endsley's Level 3 SA. Level 2 SA could be mapped to the *Knowledge* measure, or adaptation to a task's cue-criterion relations. Endsley's definition does not include the two-stage technological IP depicted in Figure 4, since it does not span the interface-mediated, human-environment system.

By inserting the sequences of technological IP and operator IP into the lens model, Stewart and Lusk (1994) developed the graphical form of the expanded lens model (ELM). The result is shown in Figure 6. Figure 6 maintains the basic configuration of the lens model, but shows the sequence of technological IP inserted into the *situation* side, and the sequence of operator IP inserted into the *operator* side. Expanding the LM in this manner introduces a new set of relationships between the model's components.

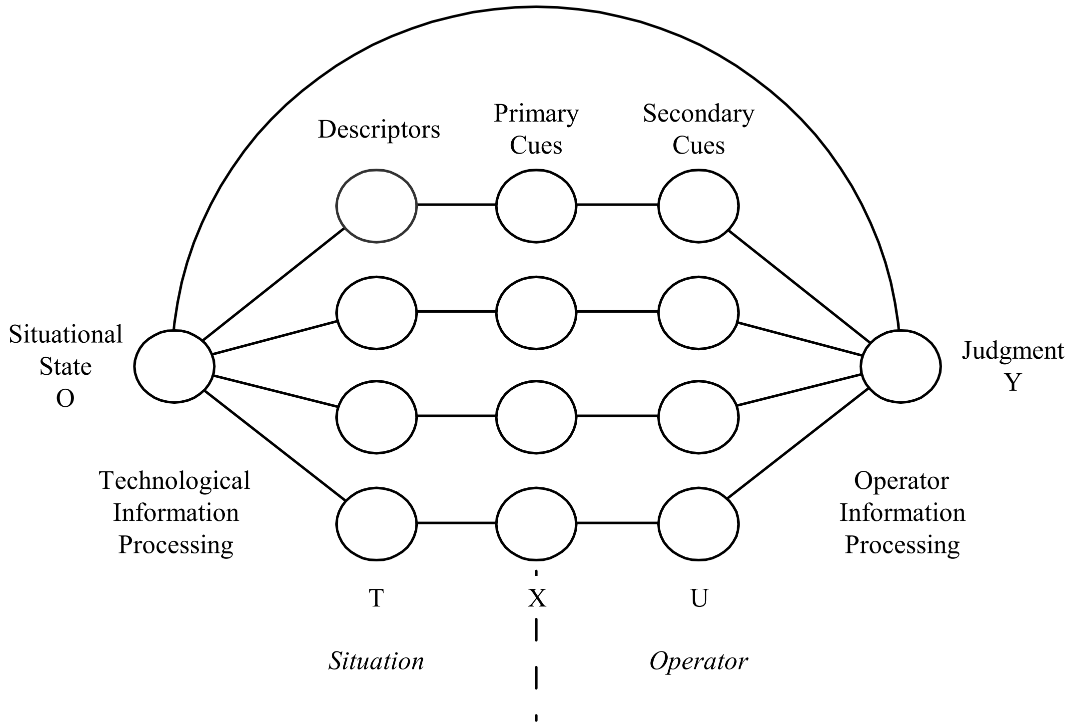


Figure 6. The expanded lens model (ELM).

For technological IP, Stewart and Lusk (1994) used correlations to measure the two relevant correspondences: (1) between the situation and the descriptors and (2) between the situation and the primary cues. Calling the correlation between the situation and primary cues $R_{O.X}$ and the correlation between the situation and the descriptors $R_{O.T}$, Stewart and Lusk derived Equation 10. This equation shows the decomposition of $R_{O.X}$ into the product of $R_{O.T}$ and the ratio of $R_{O.X}$ to

$$R_{O.X} = R_{O.T}\left( {R_{O.X}}/{R_{O.T}} \right) = R_{O.T}V_{T.X} \qquad \text{Equation 10}$$

$R_{O.T}$ (labeled $V_{T.X}$). $V_{T.X}$ is called the *Fidelity of the Information System* (Stewart and Lusk, 1994). It captures the proportion of the variance within the descriptors, relative to the situation, that is maintained in the primary cues. For example, a submarine with display technology that perfectly transformed sensor information into the cues displayed to an operator would have high fidelity (i.e., $V_{T.X}$ would equal one and thus $R_{O.X}$ would equal $R_{O.T}$). Note that these circumstances do not guarantee that the task environment is completely predictable; i.e., $R_{O.T}$ could still be less than one (e.g., the underwater environment may be inherently noisy).

$$R_{Y.X} = R_{Y.U}\left( {R_{Y.X}}/{R_{Y.U}} \right) = R_{Y.U}V_{U.X} \qquad \text{Equation 11}$$

Treating operator information processing analogously (see Equation 11), the complete mathematical form of Stewart and Lusk's expanded lens model is shown below in Equation 12:

$$\text{SS} = \left[ \overset{\boxed{1}}{(R_{O.T})} \overset{\boxed{2}}{(V_{T.X})} \overset{\boxed{3}}{(G)} \overset{\boxed{4}}{(V_{U.X})} \overset{\boxed{5}}{(R_{Y.U})} \right]^2 - \overset{\boxed{6}}{\left[ r_{YO} - \left( s_Y / s_O \right) \right]^2} - \overset{\boxed{7}}{\left[ \left( \overline{Y} - \overline{O} \right) / s_O \right]^2} \qquad \text{Equation 12}$$

Seven components comprise the final ELM decomposition. Components [6] and [7] are the regression and base-rate biases discussed previously. Components [1]-[5] result from decompositions of *Achievement*, *Environmental Predictability*, and *Consistency* from the LME. Component [3] survives from the original LME and represents the operator's adaptation to cue-criterion correlations. Components [1] and [2] capture the quality of the technological sensors that extract data from the situation ([1]), and the technological IP transforming it into the cues presented to the operator ([2]). Components [4] and [5] capture the quality of the operator IP that acquires the primary cues to form secondary cues ([4]), and transforms these secondary cues into a situational judgment ([5]). These components are summarized in Table 3.

| **ELM Component** | | **Name** |
|---|---|---|
| | SS | Skill Score |
| (1) | $R_{O.T}$ | Environmental Predictability |
| (2) | $V_{T.X}$ | Fidelity of the Information System |
| (3) | G | Knowledge |
| (4) | $V_{U.X}$ | Consistency of Information Acquisition |
| (5) | $R_{Y.U}$ | Consistency of Information Processing |
| (6) | $\left[ r_{YO} - \left( s_Y / s_O \right) \right]^2$ | Regression Bias |
| (7) | $\left[ \left( \overline{Y} - \overline{O} \right) / s_O \right]^2$ | Base-Rate Bias |

Table 3. The Components of the Expanded Lens Model

*Practical Considerations in SA Measurement*

Modeling and measuring interface-mediated SA using this approach requires that the judgment task be analyzed in accordance with the structure depicted in Figure 6: the situation must be defined as a criterion to be judged, the operator's task must be defined in terms of the judgments to be made, and information must be defined as a set of cues. Moreover, to implement the (typical) regression modeling procedure, both situation and judgment must be assigned quantitative values, and the numerical characteristics of the cues, such as their values and ranges, must be defined. This may not always be possible, or natural, and of course the model can make high demands for data collection due to the need to individually estimate its many diagnostic parameters (e.g., see Bisantz, Kirlik, Gay, Phipps, Walker and Fisk, 2001).

In addition to these limitations, the approach is not equipped to investigate all aspects of SA. For example, cognitive constructs such as memory, mental pictures, attention, and schemata (Endsley, 1995b; Wilson, 1995), are outside the purview of the ELM. Furthermore, the approach is not equipped to capture the relationships between SA and communication (Schreiber, Bell and Raspotnik, 1998), or the emotional and phenomenological aspects of SA (Gerson, 1997).

## RELATED APPROACHES

We have already discussed the relationship between our technique for modeling and measuring SA and Endsley's more cognitively-oriented theory of SA. Note that in every case we were able to portray these relationships as complementary: each approach addresses issues the other does not, and where overlap exists, we have shown that a systems-oriented approach may provide additional resources for measuring IP-related, theoretical constructs embodied in Endsley's theory. Similarly, we believe that the systems approach has a synergistic relationship with two additional, influential lines of thinking in human factors and cognitive engineering.

### *Human-Automation Interaction*

Parasuraman, Sheridan and Wickens (2000) provided a framework for human-automation interaction (HAI) by defining four "stages" of automation. Stage 1 automation concerns the acquisition of information from the task environment. Stage 2 represents a processing, fusing, or filtering of this information prior to information display. Stage 3 concerns decision support in selecting a course of action. Stage 4 concerns action implementation. For a review including a variety of applications of this framework to studies of HAI, see Wickens and Xu (2003).

Various components of our SA modeling and measuring technique map directly onto Parasuraman et al.'s framework. For example, the model decomposes overall task predictability (from the operator's perspective) into its environmental predictability and the predictability inherent in the fidelity of sensing and (automated) IP system. These measures could be used to assess, and perhaps predict, the impact of additional sensors (Stage 1 automation) or data fusion, filtering, etc. technology (Stage 2 automation). More generally, our approach to SA modeling allows one to diagnose and localize any positive or negative effects of automation on various measures (e.g., to consistency of information processing, to task knowledge, etc.) We believe that used collectively, the Parasuraman et al. framework and systems-oriented, ecological approach may advance understanding of the impact of various types of automation on SA.

### *Naturalistic Decision Making*

The Naturalistic Decision Making (NDM) paradigm (Klein, 1999), has come to represent a broadened view of judgment and decision making, with a focus on studying "how people use their experience to make decisions in field settings" (p. 97). It is crucial to note that, like NDM, the present systems approach to SA measurement and modeling does not have its basis in either classical decision theory (CDT) or behavioral decision making (BDM) approaches to the study of cognition. Both of those approaches have their foundation in models based on the internal *coherence* of cognition (e.g., expected utility theory, Bayes theorem, etc.). Like NDM, we instead use a *correspondence-based* approach, where cognition and behavior are evaluated in terms of adaptive achievement (for a discussion of the contrast between coherence- and correspondence-based approaches, see Hammond, 1999).

There are substantive ties between the two approaches as well. First, NDM acknowledges that environmental uncertainty places a ceiling on achievement. As Lipshitz, Klein, Orasanu, and Salas (2001a) have (qualitatively) put it, “Uncertainty is intimately linked with error: the greater the uncertainty, the greater the probability of making an error” (p. 339). Second, NDM, due to its historical roots in Klein’s Recognition-Primed Decision (RPD) model, places a heavy emphasis on identification of diagnostic cues (Klein, 1999) supporting inference (e.g., Crandall and Getschell-Reiter, 1993). Third, owing to the methodological roots of the present approach in the work of Brunswik (1956), our approach to the study of SA shares with NDM the goal of “conducting one’s study with representative samples of subjects. task. and contexts to which one wishes to generalize” (Lipshitz, Klein, Orasanu, and Salas, 2001b, p. 386).

Where our approach may differ from NDM, however, is that we do not necessarily agree that scientific research should include, as an “essential characteristic” a commitment to “*informal* modeling” (italics our emphasis, from Lipshitz et. al, 2001a, pp. 334-335). NDM’s commitment to informality seems to us to arise out of its cognitive rather than ecological focus: “NDM places the human . . . at the center of interest and as its basis for prescription” (Lipshitz et al., 2001a, p. 333). As such, like the work of Endsley discussed previously, NDM places a premium on the richness of its psychological constructs (e.g., “mental pictures,” “metacognition,” etc.) as these are seen to be required to paint a faithful phenomenological account of naturalistic cognition.

We recognize the value of naturalistic observation, and have made it a point to admit that our techniques do not capture all of the possible dimensions of SA. However, our view is that, while *premature* formalization is not appropriate, aiming toward increased formality should remain a guiding light in human factors, as with formality comes abstraction, the central means by which generalization can be assured, or at a minimum, can be empirically tested and evaluated.

## CONCLUSION

We have presented a systems-oriented, ecological perspective on modeling and measuring interface-mediated SA. While the approach does not speak to every theoretical construct, or deal with every phenomenon previously claimed to contribute to SA, we do believe it should be considered as an addition to the human factors toolbox, as advances in theory and application depend on advances in measurement. We have also discussed how our methods complement related approaches, not only to SA, but also to HAI and NDM. Viewed abstractly, our techniques support separating signal and noise in the performance of uncertain tasks. We appreciate that the data needed to implement these techniques may be inconsistent with the realities of field studies. Thus, we naturally do not recommend these methods for every study touching on SA issues.

We note, however, that these techniques, seen as tools for filtering signal and noise in behavior in uncertain situations, are also useful for measuring the degree to which any claims about internal psychological processes can be sustained based on sparse data sets. While the need for extensive sampling of both situations and behavior in uncertain tasks can be viewed as a legitimate limitation of our approach, this same issue should caution us about the reliability of *any* inferential technique used to identify the factors contributing to SA based on sparse data sets.

A companion article (Strauss and Kirlik, 2003) describes the first empirical test of the utility of this SA modeling and measurement approach, in diagnosing and isolating the effects of both display design differences, and also individual differences in SA achievement.

## ACKNOWLEDGEMENTS

We thank Johns Hopkins Applied Physics Lab and the Naval Air Warfare Center for support.

## REFERENCES

Adams, M. J., Tenney, Y. J., and Pew, R. W. (1995). Situation awareness and cognitive management of complex systems. *Human Factors, 37*(1), 85-104.

Bisantz, A., Kirlik, A., Gay, P., Phipps, D., Walker, N. & Fisk, A.D. (2001). Modeling and analysis of a judgment task using a lens model approach. *IEEE* Transactions *on Systems, Man and Cybernetics – Part A: Systems and Humans, Vol. 30*, No. 6.

Brehmer, B. (1994). The psychology of linear judgment models. *Acta Psy. 87*, 137-154.

Brunswik, E. (1956). *Perception and the Representative Design of Psychological Experiments*. Berkeley, CA: University of California Press.

Cooksey, R. W. (1996). *Judgment Analysis: Theory, Methods, and Applications*. San Diego, CA: Academic Press, Inc.

Crandall, B. & Getchell-Reiter, K. (1993). Critical decision method: A technique for eliciting concrete assessment indicators from the intuition of NICU nurses. *Advances in Nursing Science, 16*(1), 42-51.

Cronbach, L. J., and Gleser, G. C. (1953). Assessing similarity between profiles. *Psychological Bulletin, 50*(6), 456-473.

Durso, F. T., & Gronlund, S. (1999). Situation awareness. In Durso, F. T., Nickerson, R., Schvaneveldt, R. W., Dumais, S. T., Lindsay, D. S., & Chi, M. T. H. (Eds)., *The Handbook of Applied Cognition (284-314)*.: Wiley.

Endsley, M. R. (1995a). Measurement of situation awareness in dynamic systems. *Human Factors, 37*(1), 65-84.

Endsley, M. R. (1995b). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*(1), 32-64.

Endsley, M. R. (1996). Automation and situation awareness. In R. Parasuraman & M. Mouloua (Eds.), *Automation and Human Performance: Theory an Applications* (pp. 163-181). Mahwah, NJ: Lawrence Erlbaum Associates.

Endsley, M. R. (1997). The role of situation awareness in naturalistic decision making. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic Decision Making* (pp. 269-283). Mahwah, NJ: LEA

Endsley, M.R. and Garland, D.J. (Eds) (2001). *Situation Awareness: Analysis and Measurement*. Mahwah, NJ: Erlbaum.

Endsley, M. R., and Smolensky, M. W. (1998). Situation awareness is air traffic control: The picture. In M. W. Smolensky & E. S. Stein (Eds.), *Human Factors in Air Traffic Control*. San Diego, CA: Academic Press.

Flach, J. M. (1995). Situation awareness: Proceed with caution. *Human Factors, 37*(1), 149-157.

Gerson, C. W. (1997). Situation awareness and dynamic performance training systems: Some reflections on the literature. *Journal of Educational Technology Systems, 25*(4), 373-407.

Gibson, C. P., and Garrett, A. J. (1990). *Towards a future cockpit—The prototyping and pilot integration of the mission management aid (MMA).* Paper presented at the Situational Awareness in Aerospace Operations, Copenhagen, Denmark.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Hammond, K.R. (1999). Coherence and correspondence theories of judgment and decision making (pp. 53-65). In T. Connolly, H.R. Arkes, and K.R. Hammond (Eds.), *Judgment and Decision Making: An Interdisciplinary Reader*. New York: Cambridge University Press.

Hursch, C. J., Hammond, K. R., and Hursch, J. L. (1964). Some methodological considerations in multiple-cue probability studies. *Psychological Bulletin, 71*, 42–60.

Kirlik, A. (2000). Human factors. In Hammond, K. R. & Stewart, T. (2000). *The Essential Brunswik*. New York: Oxford University Press.

Klein, G.A. (1999). Applied decision making. In P.A. Hancock, (Ed.), *Handbook of Perception and Cognition: Human Performance and Ergonomics*, pp. 87-108. New York: Academic Press.

Lee, J.-W., and Yates, J. F. (1992). How quantity judgment changes as the number of cues increases: An analytical framework and review. *Psychological Bulletin, 112*(2), 363-377.

Lipshitz, Klein, Orasanu, and Salas (2001a). Focus article: Taking stock of naturalistic decision making. *J. Behav. Dec. Making*, 14, 331-352.

Lipshitz, Klein, Orasanu, and Salas (2001b) Rejoinder: A welcome dialogue – and the need to continue. *J. Behav. Dec. Making*, 14, 385-389.

Murphy, A. H. (1988). Skill scores based on the mean square error and their relationships to the correlation coefficient. *Monthly Weather Review, 116*, 2417-2424.

Neisser, U. (1976). *Cognition and Reality : Principles and Implications of Cognitive Psychology*. San Francisco, CA: W. H. Freeman.

Parasuraman, R., Sheridan, T.B., & Wickens, C.D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, & Cybernetics - Part A: Systems and Humans, 30*(3), 286-297.

Pew, R. W. (1995). *The state of situation measurement: Circa 1995.* Paper presented at the International Conference on Experimental Analysis and Measurement of Situation Awareness, Daytona Beach, FL.

Rothrock, L. & Kirlik, A. (in press). Inferring rule-based judgment strategies in dynamic tasks. *IEEE Systems, Man and Cybernetics – Part A: Systems and Humans*.

Salas, E., Prince, C., Baker, D. P., and Shrestha, L. (1995). Situation awareness in team performance: Implications for measurement and training. *Human Factors, 37*(1), 123-136.

Schreiber, B. T., Bell, H. H., and Raspotnik, W. B. (1998). *Investigating communication and situation awareness in air combat.* Paper presented at the Human Factors and Ergonomics Society 42nd Annual Meeting.

Smith, K., and Hancock, P. A. (1995). Situation awareness is adaptive, externally directed consciousness. *Human Factors, 37*(1), 137-148.

Stewart, T. R. (1990). A decomposition of the correlation coefficient and its use in analyzing forecasting skill. *Weather and Forecasting, 5*, 661-666.

Stewart, T. R., and Lusk, C. M. (1994). Seven components of judgmental forecasting skill: Implications for research and improvement of forecasts. *Journal of Forecasting, 13*, 579-599.

Strauss, R. and Kirlik, A. (2003). A systems perspective on situation awareness II: Experimental evaluation of a modeling & measurement technique. Ms submitted for publication.

Taylor, R. M. (1990). *Situational Awareness Rating Technique (SART): The development of a tool for aircrew systems design.* Paper presented at the Situational Awareness in Aerospace Operations, Copenhagen, Denmark.

Tucker, L. R. (1964). A suggested alternative formulation in the developments of Hursch, Hammond, and Hursch, and Hursch, Hammond & Todd. *Psychological Review, 71*(6), 528-530.

Wesler, M. M., Marshak, W. P., and Glumm, M. M. (1998). *Innovative measures of accuracy and situational awareness during landing navigation.* Paper presented at the Human Factors and Ergonomics Society 42nd Annual Meeting.