USAARL-TECH-CR--2022-17

UNITED STATES ARMY AEROMEDICAL RESEARCH LABORATORY

Aviation Decision Making and Situation Awareness Study: Decision Making Literature Review

Emilie Roth, Devorah Klein, & Katie Ernst

Notice

Qualified Requesters

Qualified requesters may obtain copies from the Defense Technical Information Center (DTIC), Fort Belvoir, Virginia 22060. Orders will be expedited if placed through the librarian or other person designated to request documents from DTIC.

Change of Address

Organizations receiving reports from the U.S. Army Aeromedical Research Laboratory on automatic mailing lists should confirm correct address when corresponding about laboratory reports.

Disposition

Destroy this document when it is no longer needed. Do not return it to the originator.

Disclaimer

The views, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other official documentation. Citation of trade names in this report does not constitute an official Department of the Army endorsement or approval of the use of such commercial items.

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
15-12-2021	Contract Report	

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Aviation Decision Making and Situation Awareness Study: Decision Making Literature Review	Prime contract No. 80ARC020D006
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
Roth, E.[1], Klein, D.[2], & Ernst, K.[3]	Letter Agreement No. AS20-01-01501
	5e. TASK NUMBER
	03
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Applied Decision Science 1776 Mentor Ave., Suite 424, MB#118 Cincinnati, OH 45212	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
U.S. Army Aeromedical Research Laboratory 6901 Farrel Rd. Fort Rucker, AL 36362	USAARL
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)
	USAARL-TECH-CR--2022-17

12. DISTRIBUTION/AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A. Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

[1]Roth Cognitive Engineering; [2]Marimo Consulting; [3]Applied Decision Science

14. ABSTRACT

We review the most prominent models of decision making from the Human Factors and Naturalistic Decision Making (NDM) research communities that are relevant to Future Vertical Lift (FVL) aviation. For each model reviewed, we briefly summarize implications for methods and measures for evaluating decision making and the impact of new technologies on individual and team decision making. The decision-making models we reviewed contribute important perspectives, methods, and measures for evaluating the effect of new technologies on decision making for both individuals and teams. Most particularly, all of the decision-making models we reviewed emphasize the importance of examining decision making under realistic conditions that reflect the challenges that arise in the real-world situations of interest. Many of the models have made methodological contributions to the design and conduct of studies evaluating decision making. Most particularly, they have emphasized the need to create study conditions (e.g., through design of evaluation scenarios) that allow important aspects of decision making to be observed and measured. Many of the models have also stimulated new measures to use in evaluating decision making. The best documented and most widely used set of new measures have come out of the literature on situation awareness, but other models of decision making have led to additional new measures as well.

15. SUBJECT TERMS

decision making, situation awareness, SA, aviation, Future Vertical Lift, FVL

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Loraine St. Onge, PhD
					19b. TELEPHONE NUMBER *(Include area code)*
UNCLAS	UNCLAS	UNCLAS	SAR	51	334-255-6906

APPLIEDDECISIONSCIENCE
1776 Mentor Ave., Suite 424, MB#118,
Cincinnati, OH 45212

Aviation Decision Making and Situation Awareness Study

Decision Making Literature Review

10 December 2021

Authors:
Emilie Roth, PhD, Roth Cognitive Engineering
Devorah Klein, PhD, Marimo Consulting
Katie Ernst, MS, Applied Decision Science

Submitted by:
Applied Decision Science, LLC
5335 Far Hills Ave, Suite #103
Kettering, OH 45429

Prepared for:
Michael Wilson, PhD
U.S. Army Aeromedical Research Laboratory

Ernie Moralez
U.S. Army Combat Capabilities Development Command
Aviation & Missile Center Technology Development Directorate

Prime contract No. 80ARC020D0006
Letter Agreement No. AS20-01501
Task 03

Executive Summary

This report is the first deliverable of the Aviation Decision Making and Situation Awareness project. The overall goal of the project is to provide recommendations for methods to systematically evaluate emerging technologies that are likely to influence or enable decision making and increase the situation awareness (SA) of Army Future Vertical Lift (FVL) aviators.

The objectives of this first report are to (1) review major theoretical approaches to characterizing decision making and SA, and (2) identify the implications of alternative theoretical perspectives for methods to operationally evaluate the impact of new technology on decision making and SA in the Army aviation environment.

For the purposes of understanding decision making in the FVL aviation environment, we use the following definition of decision making:

> *Decision making encompasses the cognitive activities involved in forming and refining a belief or course of action.*

We review the most prominent models of decision making from the Human Factors and Naturalistic Decision Making (NDM) research communities that are relevant to FVL aviation. For each model reviewed, we briefly summarize implications for methods and measures for evaluating decision making and the impact of new technologies on individual and team decision making. The models reviewed include the 'Two System' Model (Kahneman, 2011), the Recognition-Primed Decision (RPD) model (Klein, 1989) and the model of SA (Endsley, 1995). We also review the OODA Loop model, an influential model in the military that was developed by a fighter pilot (Boyd, 1987), and the Decision Ladder model (Rasmussen, 1976) that came out of the process control community as well as the family of macrocognitive models of decision making that have more recently emerged from the NDM community. We also describe two highly specialized mathematical models that have proved very useful in analyzing and evaluating the impact of new technology on human decision making – Signal Detection Theory and the LENS model.

We include a section entitled 'Pulling it all Together' that synthesizes across the set of models we reviewed to (1) identify core concepts across models that are important for characterizing decision making in the FVL environment; (2) summarize methods and measures that come out of the different decision-making modeling traditions that are relevant to evaluating the impact of new technologies on decision making in the FVL environment; and (3) present a synthesized framework of decision making relevant to FVL. This framework synthesizes core concepts that we identified to be common across the various decision-making models we reviewed and are important for modeling and supporting decision making in the FVL context.

The core concepts identified across models include:

- Decision making can arise from intuitive processes, deliberative processes, or a combination of both.

- Expert performance is often based on more intuitive, recognition-primed processes.
- Decision making is a dynamic, cyclic process, inextricable from other cognitive activities that feed into it and that it in turn influences (e.g., perception, sensemaking, planning).
- Perception involves both top-down (i.e., searching for information based on expectations) and bottom-up processes (i.e., detecting salient information in the environment that then influences comprehension and feeds further expectations).
- People actively try to understand the current situation (i.e., sensemaking) and that understanding is central to decision making.
- People develop, revise, and adapt plans as their understanding of the current situation evolves.
- Effective teamwork requires a shared understanding of the current situation and goals, sometimes called shared SA or common ground.

These core concepts provided the basis for a synthesized framework we developed to guide our next tasks.

The decision-making models we reviewed contribute important perspectives, methods, and measures for evaluating the effect of new technologies on decision making for both individuals and teams. Most particularly, all of the decision-making models we reviewed emphasize the importance of examining decision making under realistic conditions that reflect the challenges that arise in the real-world situations of interest. Many of the models have made methodological contributions to the design and conduct studies evaluating decision making. Most particularly, they have emphasized the need to create study conditions (e.g., through design of evaluation scenarios) that allow important aspects of decision making to be observed and measured. Many of the models have also stimulated new measures to use in evaluating decision making. The best documented and most widely used set of new measures have come out of the literature on SA, but other models of decision making have led to additional new measures as well. These are summarized in the report and will be explored more fully in the next phase of the project.

The synthesized framework developed in this phase of the study highlights the macrocognitive activities that enable effective decision making and how they are inter-related. It places particular emphasis on the sensemaking cognitive function (situation understanding) that generates expectations that in turn drive perception, attention, and workload management (the expectation loop). Sensemaking also generates goals that in turn drive deciding and planning as well as communicating and coordinating required for effective teamwork (the goals to action loop). The synthesized framework provides the foundation for the next set of tasks to be performed culminating in identifying methods and measures that can be used to evaluate the impact of new technologies on the various cognitive activities that underlie decision making in dynamic Army aviation.

Table of Contents

Introduction

Objectives and Scope

This report represents the first deliverable of the Aviation Decision Making and Situation Awareness project. The overall goal of the study is to provide recommendations for methods to systematically evaluate emerging technologies that are likely to influence or enable decision making and increase the situation awareness (SA) of Army Future Vertical Lift (FVL) aviators.

The objectives of this first report are to (1) review major theoretical approaches to characterizing decision making and SA, and (2) identify the implications of alternative theoretical perspectives for methods to operationally evaluate the impact of new technology on decision making and SA in the Army aviation environment.

The report begins by providing a definition of decision making relevant to complex dynamic domains such as FVL aviation. We present a real-world example of a complex military aviation decision-making task to illustrate important characteristics of decision making in complex dynamic environments such as FVL.

Next, we review prominent models of decision making from the Human Factors and Naturalistic Decision-Making research communities that are relevant to FVL aviation. In each case we briefly summarize implications for evaluation methods and measures.

We include a section entitled 'Pulling it all Together' that synthesizes across the set of models we reviewed to (1) identify core concepts across models that are important for characterizing decision making in the FVL environment; (2) summarize methods and measures that come out of the different decision-making modeling traditions that are relevant to evaluating the impact of new technologies on decision making in the FVL environment; and (3) present a synthesized framework of decision making relevant to FVL. This framework synthesizes core concepts that we identified to be common across the various decision-making models we reviewed and are important for modeling and supporting decision making in the FVL context.

The synthesized framework developed in this phase of the study highlights the cognitive activities that enable effective decision making. It provides the foundation for the next set of tasks to be performed culminating in identifying methods and measures that can be used to evaluate the impact of new technologies on the various cognitive activities that underlie decision making in dynamic Army aviation.

Definition of Decision Making

Decision making has been defined in many ways (Bisantz & Roth, 2015). For example, economists, decision-theorists, and psychologists in the judgment and decision-making tradition often define decision making narrowly as the process of choosing between two or more alternatives (Beach, 1993). From this perspective, the elements of decision making include judgment (predicting the outcomes of choosing different options); preference (weighing the costs, risks, and benefits of each outcome), and choice (combining judgments and preferences to yield a decision; Fischhoff & Broomell, 2020). Early theories employed mathematical models that assumed people are rational decision makers who compare alternatives and make the optimal choice where all the relevant constraints and preferences are known (Gonzalez & Meyer, 2016). Later theories in that tradition assumed that people fell short of rational decision makers because of systematic biases (Kahneman & Tversky, 1979; Fischhoff & Broomell, 2020).

In contrast, researchers in the Human Factors and Naturalistic Decision-Making (NDM) tradition who study decision making in dynamic, real-world contexts define decision making more broadly. In order to understand decision making in complex, dynamic contexts such as the Army FVL context, it is necessary to think about the cognitive activities that go into understanding the situation in the first place and not just the final crystallized choice. For example, in environments such as Army aviation there will be multiple disparate sources of information; the human-machine team must determine which information to attend to, and fuse information from disparate sources into a coherent understanding. There are likely to be ambiguous indicators, rapidly changing conditions, and shifting goals. This places a premium on understanding all the cognitive activities that contribute to a decision including front-end judgment processes (e.g., attending to and evaluating information in the environment and assessing the situation) as well as back-end decision processes (e.g., formulating a goal, coming up with one or more options, and making a final determination of what to do; Mosier & Fischer, 2010). Naturalistic decision-making research also suggests that decisions in complex, dynamic contexts do not necessarily involve comparison of alternatives. Rather, experienced decision-makers are likely to come up with a good solution, and then use mental simulation to imagine how it might play out, refining the solution as needed. As a consequence, for the purposes of understanding decision making in the FVL aviation environment, we will use the following broad definition of decision making:

> *Decision making encompasses the cognitive activities involved in forming and refining a belief or course of action.*

Note that the output of decision making is not necessarily an overt physical action (e.g., to navigate toward a target). It can be a decision to collect more information (e.g., to visually scan the environment to search for a potential threat) or a judgment (e.g., to identify an entity as hostile).

An Illustrative Example of Decision Making in Army Aviation

We have conducted a series of cognitive task analyses (CTAs) of Army helicopter aviation across a variety of missions including attack and reconnaissance missions (Militello et al., 2018; Militello et al., 2019; Ernst et al., 2021). This has resulted in a large corpus of cases, describing challenging decision-making situations that arose in actual Army helicopter missions. This case-base informs our understanding of the characteristics of Army aviation decision making. We provide an example here to illustrate some key characteristics of decision making in Army aviation that served to guide our review of decision-making models. It was described to us by an air mission commander (AMC) about a past incident where he flew a quick reaction force (a small team of soldiers) to a house to capture a high value individual (HVI).

> While in the air, three minutes away from the objective, the AMC observed that the house was surrounded by a compound wall – a wall that had not been present in the imagery available pre-flight. The standard procedure would be to land outside the compound where there was plenty of space for the helicopter. The AMC, however, knew that this was a time-sensitive mission. The goal was to quickly capture the HVI and leave the area before the HVI and other hostile personnel could react. He was aware that the ground force did not have ladders with them, so it would take time to either scale the wall or break through the gate. The presence of the wall introduced an unexpected and significant risk to the mission.
>
> The AMC quickly determined that he could safely land the helicopter inside the compound. Although it would require skill to land in this confined space and there was some uncertainty as he could not see everything inside the compound, he was confident he could safely land. The AMC made the recommendation to the ground force commander sitting behind him who quickly agreed. The AMC took over from the less experienced pilot flying and successfully landed inside the compound. He carefully oriented the aircraft so the soldiers would have a straight path to the door of the house, and the gunners on the aircraft could provide cover. The soldiers were able to accomplish the mission and were back on the helicopter with the HVI in minutes. (from *FVL CIS Conceptual Dynamic World Data Model* report, Ernst et al. 2021, p. 24).

In this example, the decision of interest was where to land. The challenge they faced was not the need to systematically compare alternatives. Rather what made the decision complicated was both the ‘front-end’ activity needed to understand the situation and the ‘back-end’ activity needed to convince themselves that the solution they came up with would likely work. In particular, the AMC confronted an unanticipated situation (there was a wall around the compound that had not been foreseen, the soldiers did not have ladders with them), he recognized an opportunity for action (there was enough clear space in the compound to land) that he was able to leverage to rapidly come up with a new plan adapted to the situation (land inside the compound). He simulated the plan in his mind to see if it would work (how to best orient the helicopter when it landed so that the soldiers would have a straight path to the house as they got out the helicopter door? what to do if they encountered hazards on the ground that prevented landing?) and consulted with others impacted (the ground force commander) before finalizing his decision. Finally, while the original plan was to take off immediately after dropping off the quick reaction force, the AMC rapidly revised the plan yet again because the quick reaction force was able to capture the HVI so quickly that they were able to return to the helicopter before it took off, providing a second example of revising plans to take advantage of opportunities presented. Of particular note all these intertwined cognitive activities needed to happen under extreme time pressure.

The example illustrates key tenets of decision making in Army aviation environments that

- ‘Deciding’ what to do is intrinsically tied to other aspects of cognition (such as understanding the situation including the constraints it imposes and opportunities it affords).
- Decision making does not necessarily involve concurrent comparison of multiple options.
- Decision making need not involve a ‘once and for all’ choice; often decisions need to be revised and plans adapted to deal with constraints and exploit opportunities.
- Critical decisions are often made under severe time constraints.
- Decisions often involve contributions of multiple people, especially in military contexts.

Models of Decision Making and SA

In this section we review prominent models of decision making and situation awareness that are relevant to dynamic, time-limited, high-risk decision-making environments that characterize Army helicopter missions particularly in the envisioned FVL environment.

We start with the three most prominent models of decision making and SA: The ‘Two System’ Model (Kahneman, 2011), the Recognition-Primed Decision (RPD) model (Klein, 1989) and the model of SA (Endsley, 1995). In Lipshitz’s (1993) review of different models, he specifically differentiates between models of decision making that focus on *typology* (different types of decisions) and *process* (the steps people follow). The ‘Two System’ model exemplifies the former, while the RPD and SA models illustrate the latter.

The ‘Two System’ Model

Kahneman (2011) introduced a model of reasoning and decision making that is based on the idea that people have two modes of thinking. One mode, which he calls System 1, is the intuitive mode. It operates fast, is automatic, and requires little or no conscious effort. Quickly recognizing a friend while walking down the street is an example of System 1 processing. The second mode of thinking, which he calls System 2, is the deliberate mode of thinking. It is slower, requires conscious effort, and places greater demands on memory and attention (Evans, 2008; Kahneman, 2011). Performing a mathematical calculation, making a left turn into heavy traffic, or reading a map are all examples of System 2 processing. Note that while Kahneman called these different ‘systems’ he didn’t believe that they were physically distinct mechanisms in the brain. He used the terms System 1 and System 2 as ‘fictitious characters’ intended to make the concepts being communicated more compelling and memorable.

Kahneman uses a simple example to illustrate the difference between System 1 and System 2 ways of thinking:

> A bat and ball cost $1.10.
> The bat costs one dollar more than the ball.

How much does the ball cost?

Most people quickly come up with the answer 10 cents. This includes more than 50% of students at Harvard, MIT, and Princeton who were posed this question (Kahneman, 2011). However, a little more thought reveals that 10 cents cannot be right, because if the ball costs 10 cents and the bat costs a dollar more, then the bat would cost $1.10. We know that the bat and the ball together cost $1.10 so the bat must cost less than $1.10. The right answer is that the ball costs 5 cents and the bat costs $1.05 but to come up with that answer requires more careful thought and mathematical calculation.

This example provides a compelling illustration of how System 1 and System 2 operate. System 1 generates a quick answer that intuitively feels right. In general, people will accept a plausible answer that quickly comes to mind. It is only when they decide to double check the answer using more deliberate System 2 processes that they realize their initial intuition was wrong in this case. Note that the System 2 processes required more effort and took much longer to generate the correct answer.

Note that while in this case the intuitive answer provided by System 1 happened to be wrong, it does not mean that System 1 thinking is generally error prone. On the contrary, Kahneman and others (e.g., Kahneman & Klein, 2009) strongly argue that intuition results from automatic recognition. When intuitive thought is based on expertise it tends to be both fast and accurate. At the same time, System 2 is by no means immune to error. Slower deliberative thinking can still result in mistakes (e.g., you can make a calculation error).

System 1 continuously generates assessments, feelings, and intentions. These are generally correct resulting in cognitively economical decision making without sacrificing quality. The role of System 2 is to oversee the operation of System 1. Most often this involves a low level of effort, with System 2 accepting the outputs of System 1. The exception is when System 1 does not offer a ready answer (e.g., when calculations or deliberate comparison are obviously required to generate a solution), or when some aspect of the environment suggests a need to double-check the output of System 1 (e.g., expectations are violated).

According to Kahneman's theory, System 1's intuitive judgements are derived from multiple sources. This includes expertise gained through experience. It also includes automatic mental activities related to perception and memory (e.g., quickly recognizing a face or an aircraft formation) as well as automated skills derived from extensive practice (e.g., sensory motor skills that underly driving or flying). These tend to result in rapid accurate judgments and decisions.

Another source of System 1 intuition is *heuristics.* Heuristics are cognitive shortcuts or rules of thumb that generally result in right answers but sometimes lead to systematic errors (Tversky & Kahneman, 1974).

A large literature has grown in cataloguing the different types of heuristics that people employ and the fact that these heuristics can lead to systematic bias with predictable errors. Kahneman (2011) provides a good overview of known heuristics and the biases they engender. Here we highlight some of the most important and well documented heuristics. The *availability heuristic* describes the fact that people tend to estimate how likely an event is to happen based on how easily it comes to mind. As a consequence, they tend to overestimate the probability of plane crashes relative to bus accidents because they are more likely to be covered in the news and are more memorable. Similarly, people will tend to underestimate deaths due to the flu because these deaths receive little media attention. The *representativeness heuristic* describes the fact that people judge the probability that something is a member of a class based on the degree to which it shares features with typical members of the class, often ignoring base rates. So, for example, a physician might diagnose a disease based on symptoms without considering how rarely that particular disease occurs in that specific population (also called the *base rate fallacy*). Another heuristic, *anchoring and adjustment*, suggests that people make initial assessments of a situation, and do not sufficiently update as new information is revealed. Another robust finding is the *framing effect* where objectively equivalent ways of describing a situation lead to systematically different decisions. For example, people are more likely to accept a medical procedure that is described in terms of survival rates versus objectively equivalent mortality rates (Levin, Schneider, & Gaeth, 1998). Another particularly relevant bias for the FVL aviation context is the *confirmation bias*. People have a tendency to seek out information that confirms their expectations. They are also less likely to revise their beliefs when new information comes in that is inconsistent with those beliefs (disconfirming evidence).

While the literature of heuristics and biases has focused on systematic errors that result from heuristics, there is also growing recognition that most of the time heuristics are effective in generating rapid correct answers. Gigerenzer and colleagues have repeatedly shown that simple heuristics, which they term 'fast and frugal' can lead to good performance when the heuristic is well-matched to characteristics of the task environment (Gigerenzer, Todd, & the ABC Research Group, 1999; Gigerenzer & Gaissmaier, 2011).

Kahneman's 'Two System' model and the scientific debates it has engendered have contributed to our understanding of the conditions under which intuitive judgment is likely to be valid. There is growing consensus on the conditions where intuitive judgment is likely to be valid and when it is more likely to fall short (Kahneman & Klein, 2009). For skilled intuition to develop, there needs to be valid cues in the environment that can be learned and people need to have sufficient experience and feedback to learn them. For example, Shanteau (1992) found that experts are more likely to perform well when valid cues are available and there are opportunities for feedback (e.g., insurance analysts, livestock judges) than in domains where those conditions do not hold (e.g., clinical psychologists, personnel selectors).

This research points to the need for decision aids which provide and highlight informative cues, and training programs that allow people to experience and learn valid cues. More broadly, there is a need to develop decision aids that can leverage the cognitive efficiency of System 1 while engaging System 2 when appropriate. For example, decision aids can compensate for various biases by explicitly suggesting alternative hypotheses and highlighting disconfirming as well as confirming evidence. For example, Morrison et al. (1998) describe a proposed decision-support system for evaluating the threat of unknown radar contacts, in which evidence for and evidence against a particular threat category is explicitly displayed.

Implications for Evaluation Methods and Measures

The primary implications of the 'Two System' model in design of evaluations is in the selection of test conditions and test scenarios. If the goal is to support development of expertise or to study expert decision making, then it is important to utilize conditions where there are reliable cues available in the environment that the person can sense and observe as well as reliable feedback to support development of expertise. If the goal is to evaluate the ability of a decision aid to overcome cognitive biases, then it is important to create scenarios where cognitive biases that lead to error are likely to occur. For example, test scenarios might include situations where early indications might create expectations (e.g., this aircraft is likely to be hostile) that would lead to a confirmation bias (e.g., to interpret behavior of the aircraft as hostile when in fact it was not).

The Recognition-Primed Decision Model

In reaction to models based on rational choice, the Recognition-Primed Decision model was developed based on naturalistic research on experts. Through interviews with chess grand masters and experienced firefighters, Gary Klein and colleagues (1993) found that these experts were not creating lists of decision criteria, generating options, and carefully evaluating the options against the criteria. By traditional criteria, they were not behaving rationally at all.

Instead, these experts would look at the situation and recall similar events from the past. They then evaluated the solutions they had used in those situations for how well they could work in the current context. If it seemed to work, they would proceed without considering any more options, without listing a single criterion. If, however, the prior solution failed the mental audition for this role, they would continue to find another event that seemed similar and audition that one, and so on.

Klein and colleagues (1993) argued that experts have a wealth of past experiences on which to draw and many exemplars to consider—they search for one that they recognize to prime a course of action. While criticism from a more rational choice approach seems easy –how do we know they really arrived at the best solution? —often in naturalistic settings there is no best solution to a novel problem and by the time every single potential solution has been listed out, the house has burned down, and the chess clock has run out.

In the oft-cited example of the Miracle on the Hudson, Captain Sullenberger had less than 3 minutes to come up with a plan after engine failure, during which he was also flying the plane, trying to restart the engines, communicating with ATC, and coordinating with his crew. While he certainly engaged in comparisons as part of his process (which landing sites could work?) he was not listing out all possible solutions for evaluation. Given the time pressure, it is easy to imagine how doing that could have resulted in a less satisfactory solution.

When faced with a novel, unexpected situation, one for which there was no prior referent, experts may select some near-analogous alternatives and simulate them for fit. If they can modify an alternative and make it work, they can proceed, but they may engage in more deliberative decision making and compare the leading contenders. Even in these situations, they are not being exhaustive (Klein, Wolf, Militello, & Zsambok, 1995).

Based on these observations, Klein (1989) developed a model of decision making that is called the Recognition-Primed Decision (RPD) model. The basic components of the RPD model are:

- Assessing if the situation feels familiar in any way
- If yes, assess in what respects it is familiar according to the four different aspects of recognition (familiar goals, familiar expectancies, familiar cues, and/or familiar actions).
- In particular, the person generates expectations about the situation based on their past experiences, and then assesses if these expectancies are being violated, in which case the situation must be reassessed, and more information gathered.
- If recognition is primed, the action is then mentally simulated to assess "will it work?". If yes, then implement it.
- If yes, but with some exceptions, the solution is modified and implemented.

On the other hand, if the situation does not feel familiar, the person must reassess and gather more information, as they do if expectancies are violated.

Klein (1993) recognized three variations on this basic process. These are shown in Figure 1.

Variation 1, A simple match: When factors of the decision, the context, and the goals, are familiar and typical, the decision maker can quickly move to action. With more expertise, more and more situations will seem quickly familiar and can be acted on. For novices, every situation feels new and unique.

Variation 2, Developing a course of action: When a situation is similar to previously experienced cases but differs in some respects, the decision maker may need to determine whether courses of action that have worked in the past will work in this situation. They do this by mentally simulating the actions to see if they will work for the current situation. If the answer is yes, they will select that course of action. If the answer is no, they will look for an alternative course of action and simulate it until they identify a course of action that they believe will work. Experts have an advantage of being able to quickly imagine how actions might work and either accept or reject them and moving to considering another option. Under time pressure, decision makers will usually pursue the first option they think may work. Novices have fewer options to consider and may be slower at mentally auditioning them.

Variation 3, Complex RPD strategy: When a situation is unfamiliar, the decision maker must slow down and figure out what is happening before they can move to a decision. In particular, if the expectations they generate are based on their current understanding, they will seek more information and attempt to reassess the situation. Once they are satisfied that they understand the situation they need to decide on a course of action. Even if there are only a few actions available, choosing one can be challenging. In this variation, experts mentally simulate the potential actions to see which might work better. Novices may struggle to know even how to begin.

To collect data on RPD, Klein and team developed a form of in-depth interviewing and analysis called Cognitive Task Analysis (CTA). In a CTA (Crandall, Klein, & Hoffman, 2006), the interviewer elicits a story and then unpacks the details of that story through multiple passes through, to understand the context, the cues, the mental models, and information flow of the event. The output is a deep understanding of the cognitive journey, what information was used to make decisions, and what information was never considered.

In one story collected from firefighters, a team entered a burning house to tackle what they thought was a kitchen fire. In truth, the fire had started in the basement, but the fire commander did not even know there was a basement; he just saw a kitchen filled with flames. Suddenly, he had a bad feeling and ordered his firefighters out. The kitchen floor collapsed just moments after they exited. How did he know to get them out? When interviewed, the commander did not know himself, but it made him uncomfortable that there was not a logical reason for his actions. During the CTA interview, the interviewer probed him about the details of the kitchen: the sights, smells, and sounds. In the process of the interview, the fire commander said that he realized that there was something off about the situation – although the fire was hot, it was quiet. Hot fires are normally noisy. In the moment, he was not sure what was giving him that uneasy feeling, but he rushed his team out the door because he knew something was not right. It was only in retrospect that the firefighter reflected on what was different about this fire. That quick recognition of an anomaly drove his decision making.

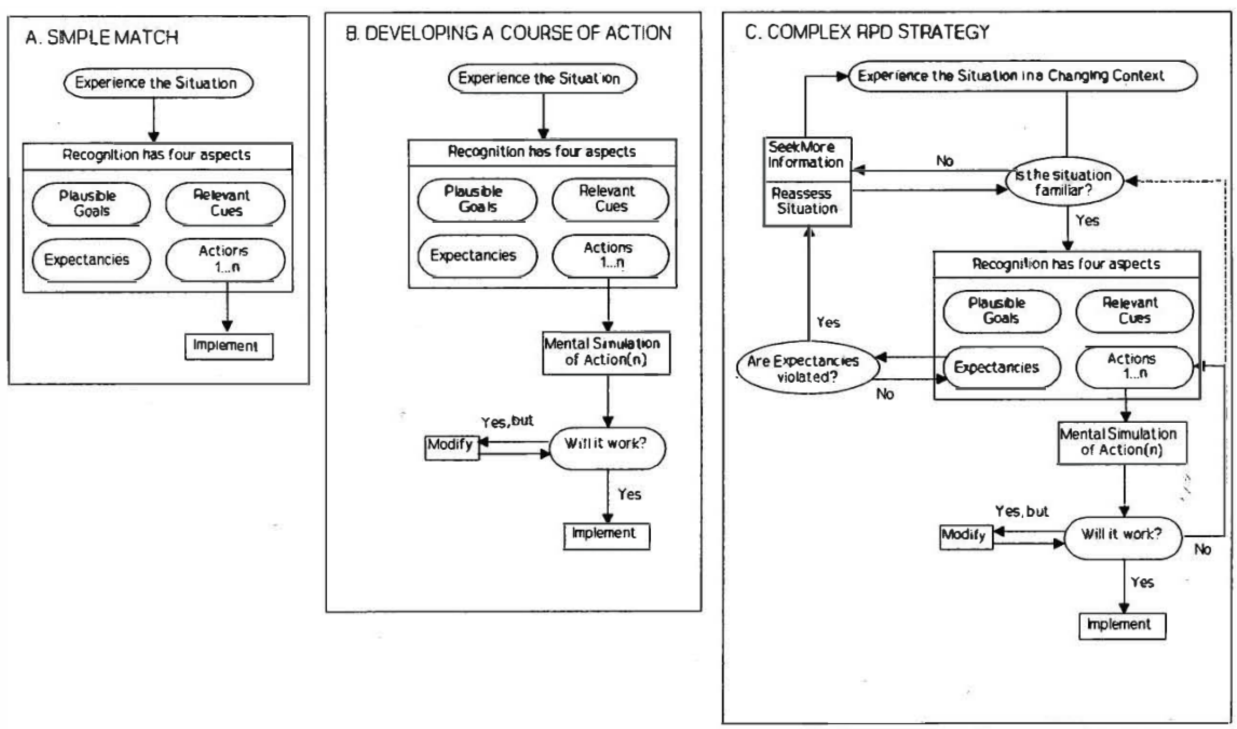

Figure 1. Three variants of the RPD model (Klein, 1993). Reprinted by permission of G. Klein.

Implications for Evaluation Methods and Measures

The methods used in development and evaluation of the RPD model are primarily naturalistic. CTA has been used to elicit stories of tough decisions and to understand the process for how they were handled. Conducting a CTA has demonstrated significant differences between experts and novices (Klein, 1993) and helped demonstrate where training can support growth of expertise.

The RPD model and the NDM community from which it arose have made a number of contributions to methods and measures for studying decision making (Militello & Anders, 2019). First and foremost has been its emphasis on studying decision making under realistic conditions, and preferably in the actual world environment. Secondly, it has provided evaluation paradigms for revealing the basis for expert decision making, including performing comparisons of decision making between experts and less experienced practitioners. In addition, CTAs are particularly useful for identifying the types of problems that are challenging (which should be used in evaluating new tools) and the contexts for them.

Militello and Klein (2013) showed how an RPD-driven approach can shape what they called Decision-centered design. By focusing on the key decision points and edge situations for evaluation, researchers can ensure that tools will be robust, meaningful, and truly helpful. They suggest using CTA interviews to elicit stories that highlight the toughest decisions and the situations that are most challenging. Using cognitive performance measures to assess how well tools aid during those moments is a targeted and precise way to see how much aid is provided in the moments that matter.

Model of Situation Awareness in Dynamic Decision Making

Among the more prominent cognitive models in Human Factors is the Endsley model of situation awareness (SA) in dynamic decision making (Endsley, 1995a; Parasuraman, Sheridan & Wickens, 2008). At its simplest, situation awareness is used to mean knowing what is going on around you (Byrne, 2015). The concept has its roots in military aviation where it was used by pilots as far back as WWI (Endsley, 1995). SA continues to be an important concept in both military and civilian aviation, although interestingly the term used in aviation tends to be 'situational awareness' (Byrne, 2015). For example, FAA guidance defines situational awareness as "continuous extraction of environmental information, integration of this information with previous knowledge to form a coherent mental picture, and the use of that picture in directing further perception and anticipating future events (U. S. Department of Transportation, Federal Aviation Administration, 2014, Watch Supervision Section, para. 11 – as referenced by Byrne, 2015).

While the concept of SA was long discussed and employed by academics and domain practitioners, Endsley provided a formal definition of SA that has since been widely adopted. She defines SA as:

The perception of the elements in 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 1995a, pg. 36).

Endsley argued that SA is critical to decision making because if SA is incomplete or incorrect, it can lead to wrong decisions. She provides as an example the U.S.S. Vincennes incident when the operators had an incorrect SA concerning an incoming aircraft that led them to believe it was a hostile aircraft. This resulted in their shooting down a commercial airliner with the loss of everyone on board (Endsley, 1995).

In her 1995 paper, Endsley presents a framework model of the role of SA in dynamic decision making. The model is shown in Figure 2.

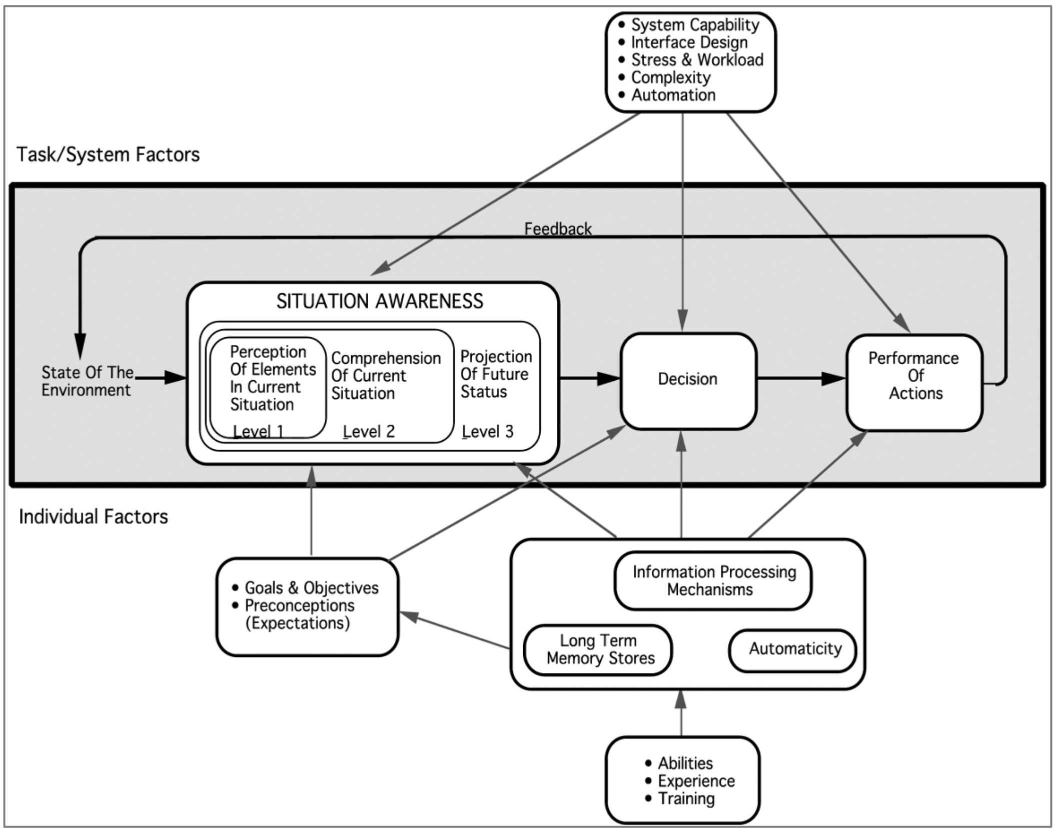

Figure 2. The Endsley (1995a) model of situation awareness in dynamic decision making. Reprinted by permission of SAGE Publications.

In this model, SA is conceptualized to include three levels:

- ***Level 1 SA - Perception of the elements in the environment:*** Level 1 SA relates to perceiving the status, attributes, and dynamics of relevant elements in the environment. For a pilot that might include information and warnings presented on displays in the cockpit, as well as aspects of the terrain and location of other aircraft.

- ***Level 2 SA - Comprehension of the current situation:*** Level 2 SA refers to understanding the significance of the perceived elements (Level 1 SA) in light of pertinent operator goals. For example, a pilot may understand that if he or she detects three enemy aircraft within a certain proximity of one another in a given geographical location, it indicates certain things about their objectives.
- ***Level 3 SA - Projection of future status:*** This refers to the ability to make near-term projections of the future state of the environment or the future actions of agents in the environment. For example, if based on Level 2 comprehension processes a pilot ascertains that a threat aircraft in a certain location is taking an offensive posture, that recognition can be used to project that the aircraft is likely to attack in a given manner.

While there has been wide attention given to Endsley's definition of SA, her discussion of the characteristics of SA and the cognitive processes that support development of SA are less widely known and understood. Since first introducing the model in 1995, Endsley has written several papers further detailing and refining the description of mental processes that influence SA (Endsley, 2000; 2015; 2021a)[1]. Some notable points include:

- The levels of SA are not intended to be thought of as a linear process first gaining Level 1 SA and then forming Level 2 SA based on Level 1 SA and then gaining Level 3 SA based on Level 2 SA. In fact, Endsley (2015a) points out that you can use Level 2 SA to infer some elements in the environment (Level 1 SA) that have not yet been observed (perhaps that there is a third aircraft hiding based on the fact that you observed two enemy aircraft that normally operate as a three-plane formation based on their doctrine).
- Levels 1, 2 and 3 SA are strongly shaped by prior knowledge stored in long term memory. Relevant knowledge includes mental models and schemas that capture generic knowledge of how physical and social systems generally work (e.g., how enemy aircraft are likely to behave based on doctrine, how an automated system in the aircraft operates, how weather systems behave). These mental models may or may not be accurate.
- SA is not a passive process of taking in information that is presented in the world. It is an active process where new information is actively sought based on current understanding of the situation (Level 2 SA), expectations generated from that understanding, near term projections (Level 3 SA), and own goals (Endsley 1995a; 2015). Current situation understanding and goals will influence where attention is directed, and how information that comes in is interpreted. For example, based on their goals and current understanding, operators may look for data to either confirm or deny their assessments or to fill in gaps (Endsley, 2015). In the Cognitive Science literature this is referred to as 'top-down' (or goal-directed) processing. At the same time, salient information (e.g., an alarm in the

[1] In 2015 there was a special issue of the *Journal of Cognitive Engineering and Decision Making* devoted to reflections and commentaries on Endsley's theory of SA. This included a paper by Mica Endsley where she argues that her theory as originally described in her 1995a paper has been misunderstood and that the purpose of her 2015 paper was to set the record straight. Several commentaries (e.g., Stanton, Salmon & Walker, 2015; Klein, 2015) have argued that her original model was not misrepresented but rather that her theory may have evolved over time. Here we describe her theory as reflected in her latest papers. We do not attempt to take a position with respect to whether the theory as described in 2015 has shifted from the theory as originally described in the 1995 paper.

cockpit) will capture attention influencing what is observed (Level 1 SA), which in turn will drive comprehension processes (Level 2 SA). This is referred to as 'bottom-up' (or data-driven) processing in the Cognitive Science literature. Thus, SA and decision making is an iterative process, with goals and understanding driving the search for new data and new data informing understanding and decisions.

- SA is not a once and final state, but is constantly being updated as the world changes, as well as the individual's understanding of the situation and/or as their goals evolve over time.
- Finally, SA will not necessarily be complete and accurate. Level 1 SA can be incomplete and/or include wrong information. Level 2 SA (comprehension of the current state of the world) can also be incomplete or wrong, and Level 3 projections can be inaccurate.

In Endsley's model, SA is a distinct stage of cognitive processing from the decision-making process and from performance of actions. She argues that while good SA (i.e., correctly understanding what is happening) is an important contributor to good decision making, you can have perfect SA but still make a poor decision or perform actions poorly (e.g., due to lack of training on appropriate procedures, misinterpretation of rules of engagement, poor tactics, lack of practice or skill). Similarly, you can have poor SA and make a correct decision (e.g., due to luck).

According to the model, goals and objectives, information processing mechanisms and limitations, abilities, long-term knowledge based on experience and training, and non-conscious mental processes that are automatic (automaticity) directly influence both SA and decision making. Similarly, system design (system capability, interface design, automation), as well as other features of the task environment (e.g., complexity, stress, and workload) will influence both SA and decision-making. This is shown via the multiple arrows in Figure 2 that point to SA, which in turn influences the decision that is made as well as via the multiple arrows that point directly to the decision. Examples of factors that directly influence decision making independent of SA include standard operating procedures, doctrine, commander's guidance, and rules of engagement.

One aspect of the Endsley model of SA that has received little attention is that she conceptualized SA to be equivalent to what some others have called a situation model (Endsley, 1995; 2000). By Endsley's definition (1995; 2000) a situation model refers to a mental model of the current state of the system (or of the world). Endsley (2000) argued that a mental model describes generic knowledge about a system. In contrast, a situation model is very dynamic and represents the person's knowledge and understanding of the *present state* of the system. In that sense, it is a specific instantiation of a more general mental model (Endsley, 2000). The distinction between a mental model and a situation model has been made in the Cognitive Science literature (e.g., Van Dijk & Kintsch, 1983) and has been used by others in the Cognitive Engineering literature to describe a constructed, continually updated, integrated mental representation of a person's current understanding of the world (Vicente, Mumaw, & Roth, 2004). Endsley (1995a) pointed out that the concept of a situation model is similar to what Rasmussen (1986) referred to as an 'internal dynamic world model.' As a concrete example, a pilot may have a mental model of the different systems in an aircraft and how they work in general. This mental model is combined with incoming data to form a situation model of the current state of the particular aircraft they are flying (e.g., current speed, fuel, engine performance). Endsley (1995a) explicitly indicated that she defined the term 'situation model' and 'situation awareness' as equivalent.

According to Endsley (1995a) the elements of SA (perception, comprehension, and projection) constitute the current situation model. These are influenced by mental models, schemata, and scripts that are stored in long term memory. Schemata (also called schemas) represent prototypical situations (prototypical states of mental models) that can enable instant comprehension of a situation via pattern recognition. Scripts of appropriate actions to take may be attached to these schemata enabling rapid decision making. Mental models, schemata, and scripts serve to direct attention to relevant aspects of the situation (influencing Level 1 SA), fill in default values when lower-level information is not available, support integration of information into a coherent understanding of the situation (Level 2 SA), and provide a mechanism for projecting future states of the system based on its current state and an understanding of its dynamics (Level 3 SA).

Goals also play an important role in shaping SA. Goals can influence which mental models are used. For example, if the goal of a pilot is to avoid obstacles in the environment, then that goal would trigger particular mental models (e.g., mental models of different types of obstacles such as wires, and where they are likely to be found in the environment) which in turn would generate expectations (e.g., wires can be spotted by looking for towers; wires are often found near roads) that would guide attention to different aspects in the environment (e.g., to spot wires look for towers especially near roads). The information obtained from the environment (Level 1 SA) would then serve to update the current situation model (e.g., there is a tower up ahead) and projections into the future (e.g., if I do not change course or altitude, I am at risk of hitting the wire). If expectations based on the current situation model are violated (e.g., I expected to come up to the tower by this point in time but have not) then the person will actively seek more information to reconcile the situation, change their mental model, and/or revise their goals and plans (Endsley, 2015).

Endsley argues that in some cases comprehension can be based on pattern matching leading to rapid decisions. This aspect of the SA model is consistent with Klein's RPD model (Endsley, 1995a). It is also consistent with Rasmussen's Decision-Ladder model (1976) that specifies that there can be short-cut links between recognizing a situation and deciding what action to take. At the same time, Endsley (2015) argues that in some cases Level 2 SA may require more deliberate sensemaking processes that involve effortfully gathering and synthesizing information using story building and mental models to develop a mental representation that accounts for and explains disparate data (Endsley, 2015).

Situation Awareness in Teams

Endsley has extended the concept of SA for an individual to teams (1995a). Endsley distinguishes two aspects of SA in teams: Team SA and Shared SA. *Team SA* refers to the degree to which every team member possesses the SA needed for his or her role. In contrast, *shared SA* refers to the extent to which team members have the same SA on shared SA requirements (Endsley and Jones, 2001).

Team SA requirements not only include information about the external situation but also about the status of teammates, their goals, and their tasks (Endsley, 2021a). This includes the status of other team member's tasks and how they might impact their own tasks, the status of own tasks and how that might impact the goals and tasks of others, the impact of own actions on others and vice versa, and projections of the actions of other team members. These elements of team SA are important to enable team members to recognize when they need to share information, coordinate action, or support the activity of others.

Shared SA reflects the idea that effective team communication, coordination, and performance requires that team members be 'on the same page' with respect to certain aspects of the situation. The concept of shared SA not only includes shared information elements (Level 1 SA) but also shared understanding of the situation (Level 2 SA) and shared projections about what is likely to happen next (Level 3 SA). The concept of shared SA is similar to the concept of common ground (Klein, Bradshaw, Feltovich, & Woods, 2005).

Implications for Evaluation Methods and Measures

One of the primary contributions of the model of SA in dynamic decision making is that it has resulted in the development of a variety of different methods for measuring SA in individuals and teams. A summary of the different SA measures and their strengths and limitations is provided in Tenney & Pew (2006) and Endsley (2021b). Measures of SA include process measures such as eye tracking, communications, and verbal protocols; performance measures such as response time and errors; subjective measures using rating scales; as well as direct measures of SA that include the SA Global Assessment Technique (SAGAT) and the Situation Present Assessment Method (SPAM).

SAGAT is one of the most widely used measures of SA (Endsley, 1995b). SAGAT has been used successfully to measure SA when evaluating system technologies, display designs, automation concepts and training programs (Endsley, 2021b).

We will review measures of SA at greater detail in a later report that focuses more directly on measures.

Decision Models From Practitioner Communities

In this section we cover two prominent models that have come out of practitioner communities: the OODA-loop model, an influential model in the military that was developed by a fighter pilot (Boyd, 1987), and the Decision-Ladder model (Rasmussen, 1976) that came out of the process control community and is widely used to support analysis and design of decision aids. These models are included in this review because they have striking similarities to models of decision making and SA that have come out of the NDM and human factors communities, reinforcing the key tenets that underlie decision making in dynamic, high-risk contexts such as Army FVL.

The OODA Loop Model

In the 1970's, a military strategist and fighter pilot named John Boyd developed the OODA loop to help support faster decision making under pressure (Boyd, 1987). The OODA loop was developed in parallel to and somewhat independently from research in cognitive and human factors psychology, but they have much in common. In the OODA loop, people are continuously cycling through:

Observing: Collecting information about the world and what is happening. As a fighter pilot, Boyd was particularly interested in how to help aviators get more information faster to enrich the observation process.

Orienting: Analyzing and synthesizing information into what he called a "mental perspective" which aligns to comprehending the situation (Endsley, 1995a).

Decision: The output of Orienting is a decision about what to do next.

Action: Putting the planned Decision into action. Again, as a pilot, Boyd wanted to increase the speed and ease of each step, so some of his thinking about Action is about how to make controls easier to use and more responsive.

Boyd saw the loop as a continuous cycle, as actions result in new stimuli to observe and incorporate (Fadok, 1995). For Boyd, the crux of winning versus losing depended on the rapidity of moving through the OODA loop. He viewed the winner as the one who could repeatedly move through the cycle faster than the enemy, so that eventually the enemy's reactions are totally inappropriate to the situation. He viewed rapid and effective orientation as the most important element.

Boyd (1976) created several diagrams depicting his model, including a high-level version that emphasizes the cyclic nature of decision making (see Figure 3) and a more elaborated version that provides more details and shows linkages across the different cognitive activities (see Figure 4). The high-level version is easy to comprehend and accessible, which certainly helped popularize it.

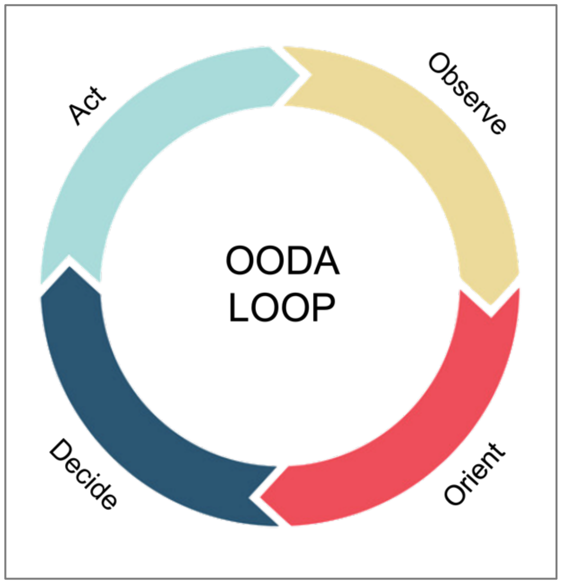

Figure 3. The simplest version of the OODA loop model stresses the cyclic nature of decision making.

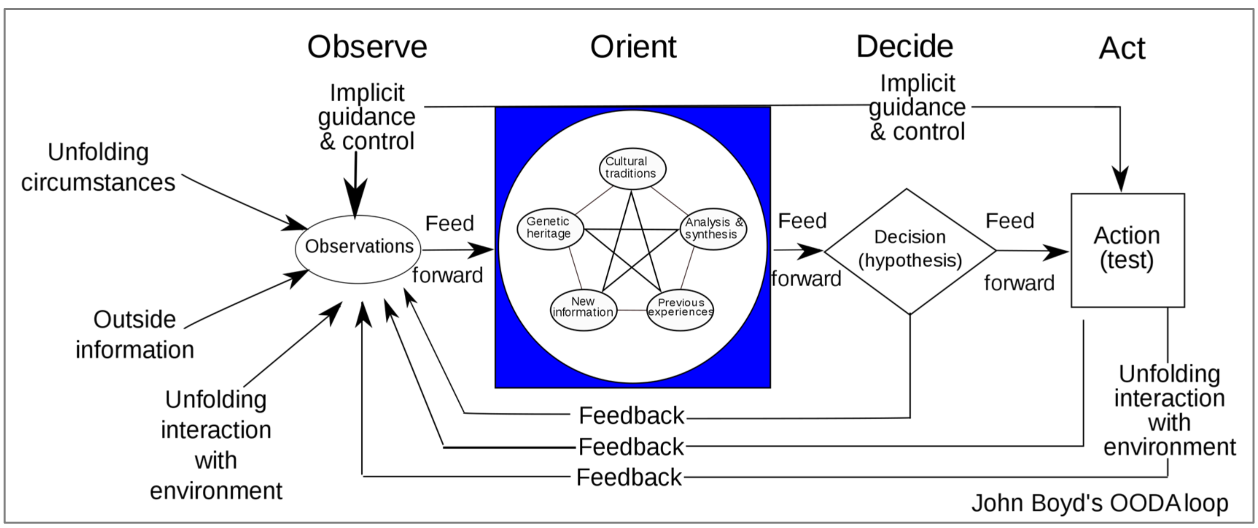

Figure 4. The more elaborate version of the OODA loop model. Original drawing by Patrick Edwin Moran obtained from https://en.wikipedia.org/wiki/OODA_loop (CC by 3.0).

In its more elaborated version, Boyd provided more details with respect to the content and inter-relationships among the elements of the OODA loop:

- *Observe*: Collecting sensory inputs from multiple aspects of the environment. This is an aggregation step, getting as much information as possible without analyzing it. Boyd felt that anything that increased the ability of a person to collect high quality inputs improved the quality of a decision. Boyd considered physiological abilities and the environment as key determinants of observational quality.

- *Orient*: Creating a mental image of reality based on the inputs through "destruction and creation" (Boyd, 1976). He argued that people build a mental image of the situation, but then destroy, or revise it, as new information is learned. The concept of a mental image is similar to what others have referred to as a mental model or a situation model. He argued that the mental image that is constructed is shaped by a person's background and experiences. Like the RPD model, Boyd believed that previous experiences play a very important role in orienting. He argued that observations that match up with past experiences call for certain decisions and actions allowing for more rapid response. In contrast, a mismatch between the real world and the person's mental image of the world will generate confusion and/or inaccurate responses. Boyd saw orient as the most important step of the loop since it shapes the success for the rest of the loop. Its impact on observation is bidirectional: shaping how we observe as well as being shaped by observation (discussed by Angerman, 2004).
- *Decide*: As the mental image evolves, it can be used as the basis for making decisions. The better the quality of orientation, the better decisions can be made. While multiple options are considered, orientation has pointed towards the best decision given what is known.
- *Act*: Once a decision is made, it is acted on. Beyond action though, this step includes evaluation of the action to ascertain its success.

Criticism of the OODA loop focuses on its testability and scope (Brehmer, 2005). There has not been much evidence collected linking the use of the loop to improved performance. In addition, OODA was developed around tactical decision making, for fighter pilots. Strategic decision making is certainly more multi-layered, collaborative, and intense, and the loop does not apply as well.

Criticisms aside, the OODA loop has been incredibly valuable to the military, who focus training and development of equipment on moving through the loop with greater proficiency. Some of the impact was on the design of equipment – for example to improve the visibility of targets from a cockpit—but Boyd's work also has influenced cognitive training programs.

While the original OODA loop was conceptualized as an individual process (in the context of dogfighting), variants such as M-OODA (modified OODA; Rousseau & Breton, 2004) and DOODA (Dynamic OODA; Brehmer, 2005) have been developed to incorporate team decision making and more strategic planning functions.

Implications for Evaluation Methods and Measures

One of the most important learnings of this approach is that to make better, faster decisions, it is important to improve the quality of the information and processes that lead to them. This is critical in the task of determining how to assess new decision-making tools for aviators in the FVL context. If we accept that the cognitive activities leading up to a choice or judgement determine decision quality, we must focus on assessing how tools are aiding along the way. As Boyd might say—all the tools in the world to help weigh different options are meaningless if you cannot see out the windshield to what the enemy is doing.

One example of how the OODA loop was applied was in an analysis Boyd did of the F-86. While the aircraft was inferior to the MiG-15s flown by North Korea in many measures of performance, the F-86 had two critical features that could help explain U.S. superiority. One was the design of the canopy—the bubble shape helped pilots see more, improving their observation. A second was the hydraulic controls which gave U.S. pilots an edge when they acted, with faster reaction time (discussed by Angerman, 2004). What does this mean for evaluation? It is important to consider not just the cognitive middle of decision making, but the sensory and physical ends of the process as well. If it is hard to collect data, if it is hard to implement a decision, the strength of the decision analysis itself is irrelevant.

The Decision Ladder Model

One prominent decision-making model that comes out of the process control industry is the Decision Ladder representation developed by Jens Rasmussen (Rasmussen, 1976; Rasmussen, Petjersen & Goodstein, 1994)[2]. The model was motivated by a set of studies that he performed examining the problem solving and decision making of process control operators as they started up a thermal power plant as well as that of maintenance technicians during equipment troubleshooting tasks. Rasmussen used a verbal protocol approach to document what the operators and maintenance technicians verbalized as they performed their work.

Rasmussen observed that in most situations the workers seemed to know what was going on and where to focus their attention without much deliberative thought. He noticed that operators tended to verbalize their current state of knowledge (e.g., what they observed, what they concluded, what they decided) but not the specific reasoning they used in coming up with those judgments. Further, they tended to express their state of knowledge in terms of assessments rather than the individual data pieces that led to those assessments. They did not mention specific instrument readings but rather talked about the current state of the plant and how they expected the plant state to change as they performed control actions. It was only when operators faced an unfamiliar situation (e.g., when there was an unexpected plant malfunction) that they started to describe their observations and reasoning process at a more fine-grained level.

From these sets of observations, Rasmussen concluded that people operate using different mechanisms, sometimes applying deliberative problem solving but more often responding in a more intuitive manner. He formulated a model of decision making that is composed of a sequence of mental processing steps and the associated knowledge state that result from each step. The full sequence of steps constitutes a complete deliberative decision-making process and is shown in Figure 5. It is called the Decision Ladder because of the shape of the sequence of steps as he represented them. A key feature of the Decision Ladder model of decision making is

---

[2] There have been questions raised whether the Decision Ladder should be considered a model or a template (Lintern, 2010). Rasmussen (1976) himself referred to it as a 'schematic map' with the idea that it represented the mental activities and knowledge states that could be used, rather than ones that are necessarily used. However, following Lintern (2010) we have included it as a decision-making model in that it provides a description of how decisions are made.

that people will rarely go through every step sequentially. Rather steps will be skipped and shortcuts through the decision ladder will be made depending on the level of skill of the individual in the specific task. He stressed that experienced workers will rarely follow the decision ladder in linear sequence. The linear sequence will only be followed when heuristic or rule-based shortcuts developed through experience are unavailable, such as when an expert is confronted with an unfamiliar task or when a novice is performing the task. Even in those cases where a more linear, methodical sequence is followed, there is no presumption that every step in the ladder will be followed.

Figure 5 depicts the detailed steps in the Decision Ladder. The rectangular shapes represent mental processes such as the perceptual system being activated, cues being observed, and situation states being identified. The circles represent the current state of knowledge that results from the mental processes. Examples of states of knowledge include being consciously aware of a set of observations, understanding what state the system is in, deciding on a goal to pursue, as well as deciding what procedure to use. As can be seen in Figure 5, the steps on the left side of the Decision Ladder relate to situation assessment and understanding the situation. The top center portion of the Decision Ladder relates to predicting the consequences of the situation (interpret), identifying options in terms of goals to pursue, and selecting a goal. The right side of the Ladder is about formulating and executing a plan for accomplishing the chosen goal.

As mentioned earlier, these steps are not assumed to be always followed in rigid sequential order. On the contrary, depending on level of expertise and familiarity with the situation, shortcuts across the Decision Ladder can be taken that skip entire sets of steps. For example, an individual can immediately recognize a system state (e.g., via pattern matching) and from that knowledge immediately know what needs to be done and which procedure to follow. In particular, Rasmussen indicated that "holistic perception leads to observation directly in terms of system state or task to perform rather than observation of separate items of information" (Rasmussen, 1976, pg. 5). In some cases, the connections between perception and action can happen in a completely automatic manner such as pressing the brake in a car when you see the brake lights of the car in front of you come on.

Rasmussen defined three modes of decision making that he termed 'skill-based,' 'rule-based,' and 'knowledge-based.' Skill-based refers to highly automated performance such as the sensory-motor control actions associated with driving a car. Rule-based refers to simple rules or heuristics that govern decisions such as slowing down and coming to a stop when you see a stop sign. Finally, knowledge-based performance is the deliberative reasoning process that occurs when faced with an unfamiliar situation. An example is confronting an unexpected street closure while driving might trigger knowledge-based assessment of your current location with respect to your goal and require a new route to be planned.

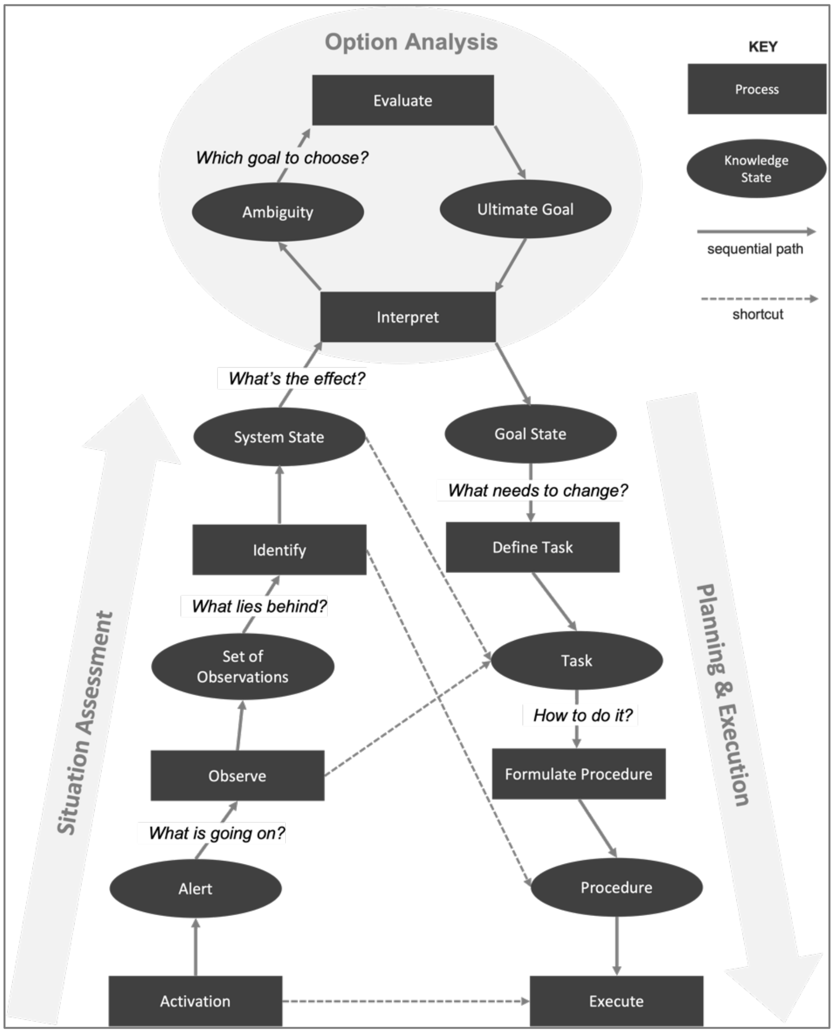

Figure 5. The Decision Ladder model depicts the processes and knowledge states associated with different steps in the decision-making process (Rasmussen, 1976). Importantly, shortcuts allow experts to skip across multiple steps of the decision ladder when making decisions in familiar situations.

Rasmussen's Decision Ladder model has strong similarity to Kahneman's 'Two System' model (Kahneman, 2011) in that it accommodates both deliberative and heuristic decision processes. Rasmussen argued that the two modes of operation are complementary, with the more automated processes enabling successful response in situations that demand rapid reaction and the more deliberative processes enabling successful response in novel situations that require reasoning from first principles, projecting consequences, and weighing alternatives.

Rasmussen's Decision Ladder model is also consistent with Klein's RPD model (1989). As both Lintern (2010) and Naikar (2010) have noted, Rasmussen's model explicitly accommodates recognition-primed decision making through the various short-cuts that are available for traversing the Decision Ladder. In particular, Rasmussen recognized that an individual could immediately perceive a situation based on pattern-recognition and immediately know what to do based on past experience. Naikar (2010) showed that other variants of decision making identified by Klein (1989) such as ones that require diagnosing a situation and/or evaluating a course of action could also be accommodated within the Decision Ladder model via different patterns of short-cuts.

Implications for Evaluation Methods and Measures

The Decision Ladder model has primarily been used to support design of decision aids. Most commonly it is used to document the cognitive tasks associated with challenging decisions in different domains. This provides the basis for specifying information needs and design requirements for new forms of cognitive support such as new displays or new forms of automation. A review of how the Decision Ladder has been used to support design is found in Roth, Sushereba, & Diiulio (2018).

The Decision Ladder can also be used to support design of studies to evaluate the impact of new forms of decision support or automation. For example, the Decision Ladder can be used as a template to trace how a new form of automation or decision support impacts the human decision-making process. Process trace methods, including eye movement measures and think aloud methods can be used to assess whether and how the introduction of a new system changes the path taken through the Decision Ladder. For example, does the new system enable the individual to more rapidly assess the situation and know what to do (e.g., via pattern recognition)? Does it effectively support performance in unanticipated situations when the person needs to use first-principle reasoning to understand what is going on and come up with a novel course of action? The Decision Ladder can also be used to guide selection of evaluation scenarios. For example, if the new system is intended to foster skill and rule-based decision making, including recognition-primed decisions, then the evaluation study should include familiar situations where recognition-primed decision making is expected. In contrast (or in addition), if the new system is intended to support knowledge-based reasoning in unanticipated conditions then the set of evaluation scenarios should include complex conditions that are likely to be unfamiliar.

Macrocognition

The Macrocognition approach (D. Klein, H. Klein, & G. Klein, 2000; G. Klein, et al., 2003, and see Vicente, Mumaw, & Roth, 2004 for related thinking) is a more recent effort to develop a naturalistic framework of decision making. The term "macrocognition" was originally coined by Pietro Cacciabue and Erik Hollnagel (1995). Macrocognitive functions refer to individual and team cognitive functions that are performed in natural decision-making settings (e.g., detection, sensemaking, planning). Macrocognitive models characterize individual and team cognition at a functional level of description that is most relevant to performance in natural settings (Klein et al., 2003; Klein, Moon, & Hoffman, 2006; Patterson & Miller, 2010). By functional level, we mean specifying the cognitive functions that underlie decision making without describing how those cognitive functions are performed at the detailed cognitive-processing level. This contrasts with microcognitive-level descriptions that tend to characterize cognitive processes at a more detailed, millisecond time scale, information-processing level (Klein, Klein, & Klein, 2000).

Multiple models describing how Macrocognition works have been proposed. We will review a few of the more widely disseminated ones, as a strong model can help inform how we think about which measures are most relevant in the FVL environment. This approach is, like the OODA loop, grounded in decision making as an iterative cycle. One important feature of Macrocognition is that it attempts to encompass both individual as well as team decision making, so concepts such as sensemaking should be considered both as an individual's sense of the world and the team's shared understanding which is grounded in communication and sharing of knowledge.

Klein et al. (2003) were among the first to call out important macrocognitive tasks that arise in dynamic naturalistic decision-making contexts, including military aviation. Figure 6 provides an overview of the macrocognitive tasks captured by their model. The macrocognitive functions listed horizontally are viewed as key macrocognitive functions that emerge in real-world settings. The macrocognitive functions around the circle are supporting functions that serve as a means to achieve the key cognitive functions (Schraagen, Klein, & Hoffman, 2008). Both the key macrocognitive functions and the supporting functions play important roles in decision-making in dynamic, high-risk environments such as military aviation.

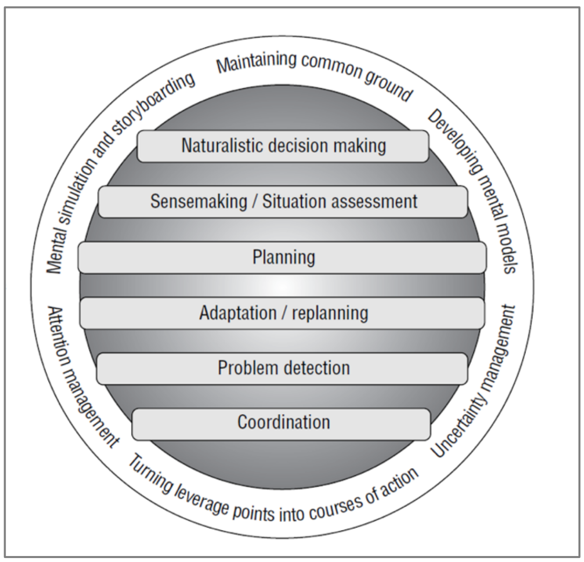

Figure 6. The Klein et al. (2003) model of macrocognition. Reprinted by permission of G. Klein.

A core macrocognitive function is *sensemaking*. Sensemaking can be thought of as how people make sense of their experience in the world (Klein, Moon, & Hoffman, 2006). It is a concept related to the notion of comprehension in Mica Endsley's SA model (Level 2 SA) in that it relates to the ability to connect the various information that have been gathered into a meaningful description of the situation. Sensemaking refers to the cognitive activity of creating an understanding, whereas Level 2 SA refers to the product of that cognitive activity. Mica Endsley used the term situation assessment to refer to the sensemaking macrocognitive function.

Klein and colleagues presented a Data-Frame theory of sensemaking (Klein, G., Moon, & Hoffman, 2006; Klein, G., Phillips, Rall, & Peluso, 2007). The main elements of this theory are presented in Figure 7.

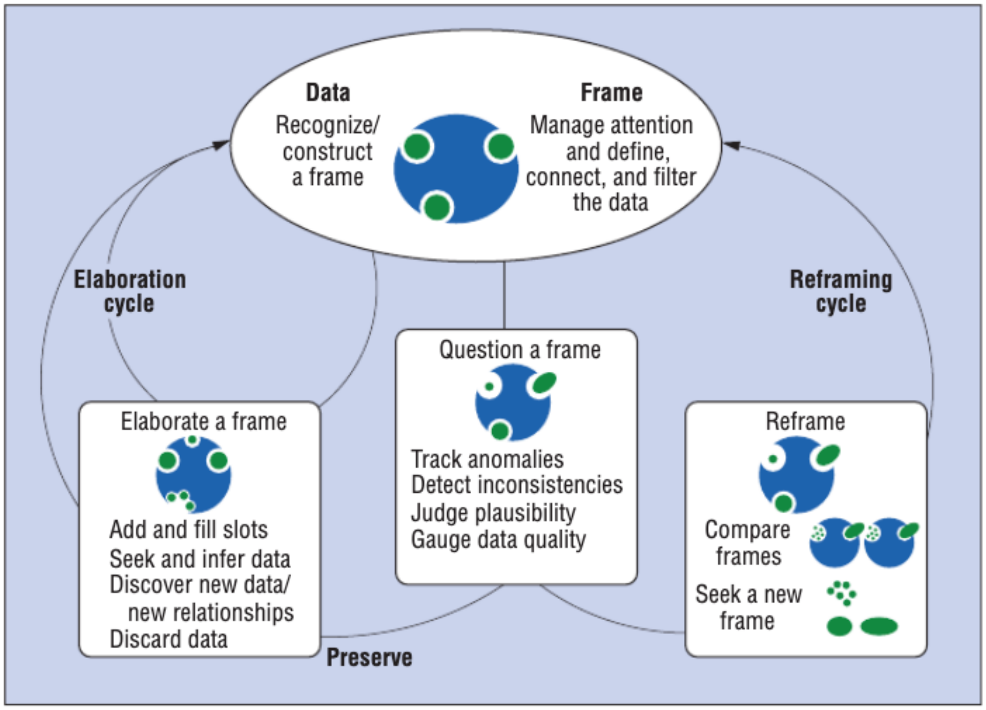

Figure 7. The Data-Frame theory of sensemaking (Klein et al., 2006). Reprinted by permission of G. Klein.

In this model, sensemaking happens through the use of 'frames.' A frame can be thought of as a perspective or frame of reference that is used to interpret and integrate incoming data. Examples include a mental map of a geographic area, a mental model of how a physical system works, or an understanding of enemy doctrine. When new data come in, the data are used to construct a frame if new or stimulate the person to recognize a pre-existing frame. New data can also help the person elaborate on an existing frame and refine what data are included in it. Inconsistencies and anomalies may force reconsideration or questioning of the current frame, which may result in reframing how information connects. For example, a military pilot observing individuals on the ground might initially interpret their behavior as benign, but that interpretation may be completely revised if the pilot notices that one of the individuals is wearing a non-standard uniform.

In the Data-Frame theory, the individual's frame guides what is considered relevant data, and at the same time observed data are integrated and used to refine the frame. As the authors state, "Sensemaking is the process of fitting data into a frame and fitting a frame around the data." A person's goals also drive how the sensemaking process occurs.

It is important to note that sensemaking is also a forward-looking task. As Endsley discusses (1995a), part of situation awareness (as she refers to sensemaking) is projecting what might happen in the future, which can guide both attention and help shape planning and decision making.

Other important functions identified in the Klein et al. (2003) model of Macrocognition that were not previously emphasized are adaptation and replanning, turning leverage points into courses of action, and uncertainty management. The *Adaptation and replanning* function emphasizes that in dynamic high risk worlds 'the plan never survives contact with the enemy.' There is a constant need to revise one's understanding and adapt one's response to the confronted situation. Similarly turning leverage points into courses of action highlights the ability of experts to exploit opportunities presented. This is a point that was also highlighted by Boyd's OODA Loop model.

G. Klein and colleagues (2007) pointed out the importance of *uncertainty management* and the challenges of working in a situation where there is missing information. Though it may be tempting to ignore what is unknown, there is a cost and expert decision makers manage uncertainty in an active way. Klein identified tactics that experts use to do this, including delaying until more information is known, increasing attention to learn more, and filling gaps with assumptions, among other options.

Klein et al. (2003) also highlighted macrocognitive functions that are important for developing and maintaining effective team performance, including the performance of military flight crews. These include coordination and managing common ground. *Common ground* refers to shared knowledge, beliefs, and assumptions that support interdependent action across a team (Clark & Brennan, 1991; Klein, Feltovich, Bradshaw, & Woods, 2005). Common ground includes having a shared understanding of the current situation as well as shared understanding of the team's goals and individual member's roles in achieving them. Common ground allows team members to anticipate each other's needs and actions, and back each other up when needed. It also allows team members to coordinate without explicit and lengthy communication (Klein, et al., 2005).

While sensemaking is generally included as a core macrocognitive function in all the models, later models have tended to include slightly different lists of macrocognitive functions depending on their focus. For example, Patterson and Hoffman (2012) narrowed the list down to four core macrocognitive functions: sensemaking, detecting problems, deciding, and replanning. They highlighted that these macrocognitive functions are core to decision making in dynamic, high-risk environments, and that they are intrinsically cyclic and intertwined. Problem detection may trigger sensemaking and in turn sensemaking may trigger a change in goals and adaptive replanning. Patterson and Hoffman used the term replanning rather than planning to emphasize that in dynamic, high-risk worlds people are always in a state of revising and adapting plans.

Hoffman and Yates (2005) focused on the fact that decision making is inherently a more fluid and comprehensive process than is captured by the concept of a discrete decision-making event. Elements include uncovering and revising goals and values as events evolve; generating and recognizing options as well as changes in options (both closed doors and new opportunities) as events shift; projecting consequences of options; and generating novel options that simultaneously meet multiple objectives, obviating the need for goal-tradeoffs.

Researchers have continued to define and focus on different macrocognitive functions, reflecting the fact that macrocognition is intended to encompass a broad range of high-level cognitive tasks that arise in real-world context. For example, Klein and Wright (2016) presented a list of macrocognitive functions that include complex learning, managing risk, and insight in addition to the more traditional set of macrocognitive functions.

Roth and colleagues (Roth, Mosleh, Chang, Richards, Bley, Shen, & Zoulis, 2012; Chang, Bley, Criscione, Kirwan, Mosleh, Madary, Nowell, Richards, Roth, Sieben, & Zoulis, 2014) developed a macrocognitive model that differs from others in that it is more explicit in presenting the linkages between different macrocognitive functions and how they collectively contribute to decisions in dynamic, high-risk worlds. This model is depicted in Figure 8.

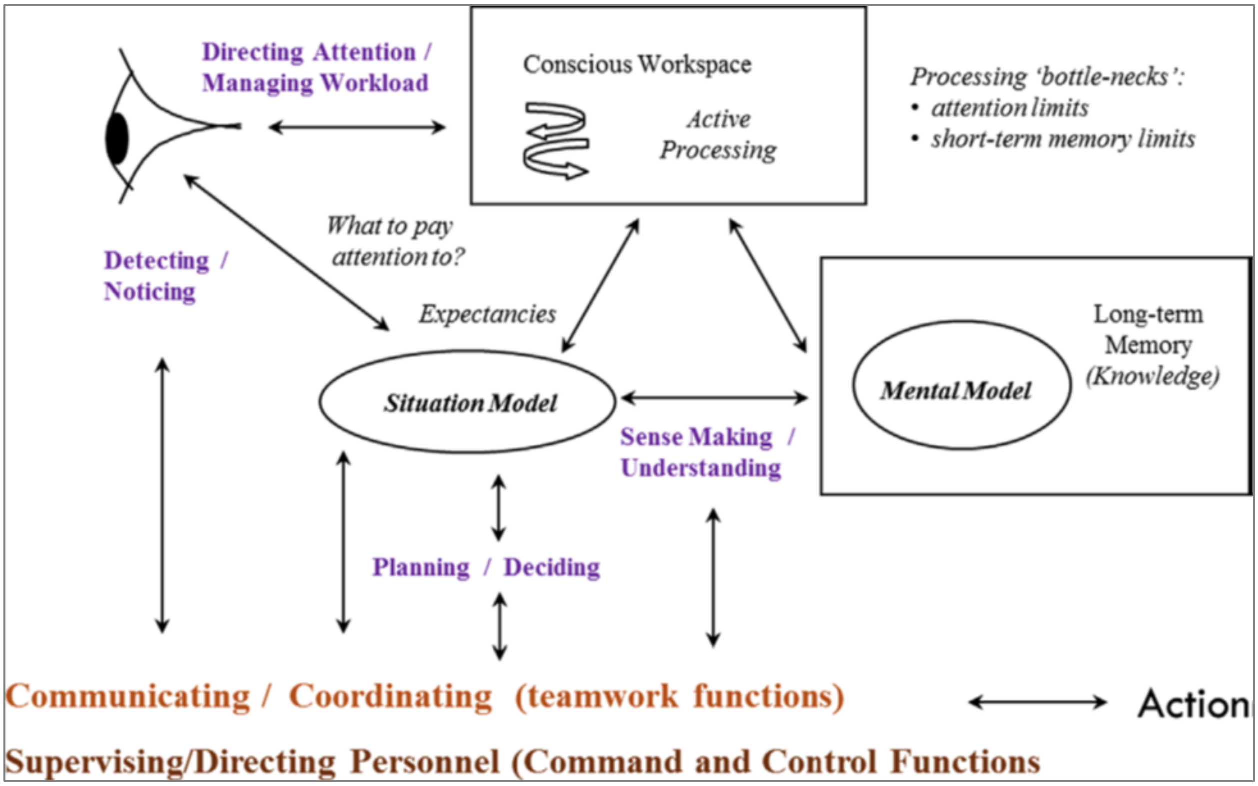

Figure 8. The Roth et al. (2012) model of macrocognition. Reprinted by permission of E. Roth.

At the core of the Roth et al.'s (2012) model is the concept that people actively work to construct a coherent understanding of the situation they are in ('*sensemaking*'). The output of sensemaking is a *situation model* that represents a person's understanding of a situation. Similar to Endsley's model of SA, the situation model draws on both real-time information obtained from the world via perceptual processes (i.e., *'detecting/noticing'*), as well as background knowledge stored in long term memory (e.g., mental models). A person's situation model may or may not be an accurate representation of the true state of the world. Another core concept of the model is that people have limited processing resources (e.g., attention, working memory). This places a premium on '*directing attention and managing workload'* functions. Attention/workload management refers to determining where to direct attention and focus activity under high workload/high attention demand conditions. People form expectations as to what should happen next and what is highest priority to deal with based on their situation model. These expectations influence where people will direct their attention and how they will manage their workload (e.g., how they will prioritize activities under high workload conditions). These in turn will influence what people will pay attention to and therefore what they will detect/notice (Mumaw, Roth, Vicente, & Burns, 2000; Vicente, Mumaw, & Roth, 2004). People's understanding of a situation also influences *planning and deciding* functions. Based on their situation model people will prioritize goals, make decisions, and plan actions.

Communicating and coordinating and other related teamwork activities, including supervising/directing personnel, are also central macrocognitive functions. These functions enable the team to operate as a cohesive macrocognitive unit. The output of all these macrocognitive processes is the execution of observable physical actions.

To summarize, the macrocognitive perspective offers a number of important insights that advance our understanding of decision making in naturalistic contexts and how to more effectively support decision making through training and decision aiding. First, it highlights that macrocognitive functions are inherently intertwined rather than sequential stages. Second, it highlights the constructive nature of sensemaking that enables people to develop and update their understanding of evolving situations. Third, it highlights that planning and decision-making processes are not 'one-off' processes that occur at a single point in time, but rather unfold dynamically as situations evolve over time and goals shift (Klein, 2007 a & b; Hoffman & Yates, 2005). As Klein (2007b) has argued, discovering and refining goals while pursuing them is at the core of adaptive, resilient performance. Finally, the macrocognitive perspective acknowledges that decision making is not limited to a single individual. In most cases decision making involves communication and coordination among multiple people that need to share a common understanding of the current situation and goals (i.e., common ground).

Implications for Evaluation Methods and Measures

When we start to consider the implications of the Macrocognition approach for evaluation measures and metrics, it is helpful to think about the different components distinctly. Patterson and Miller's 2010 book entitled *Macrocognition Metrics and Scenarios* has a wealth of information on this. Some to note:

- It is particularly important to create a strong scenario for evaluation (see Patterson, Roth, & Woods, 2010).
- There are many specific tests to probe the cognitive processes that make up the Macrocognition model (for discussion, see Klein, 2010).
- Rather than focus on particular values on a scale or the classic HEAT index (Hits, Errors, Accuracy, and Time), it can be more meaningful to look at changes over time. (Hoffman, 2010).
- Indicators for assessing systems from a macrocognitive perspective have been proposed (Wiggins and Cox 2010).

Mathematical Models of Decision Making

We end the review of decision-making models with two highly specialized mathematical models that have proved very useful in analyzing and evaluating the impact of new technology on human decision making – Signal Detection Theory and the LENS model. Unlike the models reviewed above that attempt to encompass a broad set of cognitive activities that underlie decision making, these two models focus more narrowly on the 'front-end' aspects of decision making related to judging a situation (e.g., deciding whether an entity is a threat).

Signal Detection Theory

Signal Detection Theory (SDT) is a specialized mathematical model that has proved very useful in analyzing decision making in situations where an individual (or automated system) needs to decide between two mutually exclusive states of the world based on a set of indications (Green & Swets, 1966). Real world examples include a screening agent at an airport who needs to decide whether there is a weapon in a piece of luggage based on an x-ray image; a military pilot who needs to decide whether a signal on a cockpit display indicates a threat; or a radiologist deciding whether a tumor is present based on reviewing an x-ray.

SDT conceptualizes the decision task as one of attempting to discriminate a 'signal' (the target of interest) from background 'noise.' This formulation of the problem comes from thinking about situations such as detecting a threat using sonar against general background 'noise.' The conceptual model is shown in Figure 9. The assumption is that there are two overlapping distributions: one for the case where there is a signal and one for the case where there is no signal. The more overlap between the distributions, the more challenging the decision task. As shown in Figure 9 the discrimination is more challenging in Case 1 than in Case 2 because in Case 1 the two distributions overlap much more.

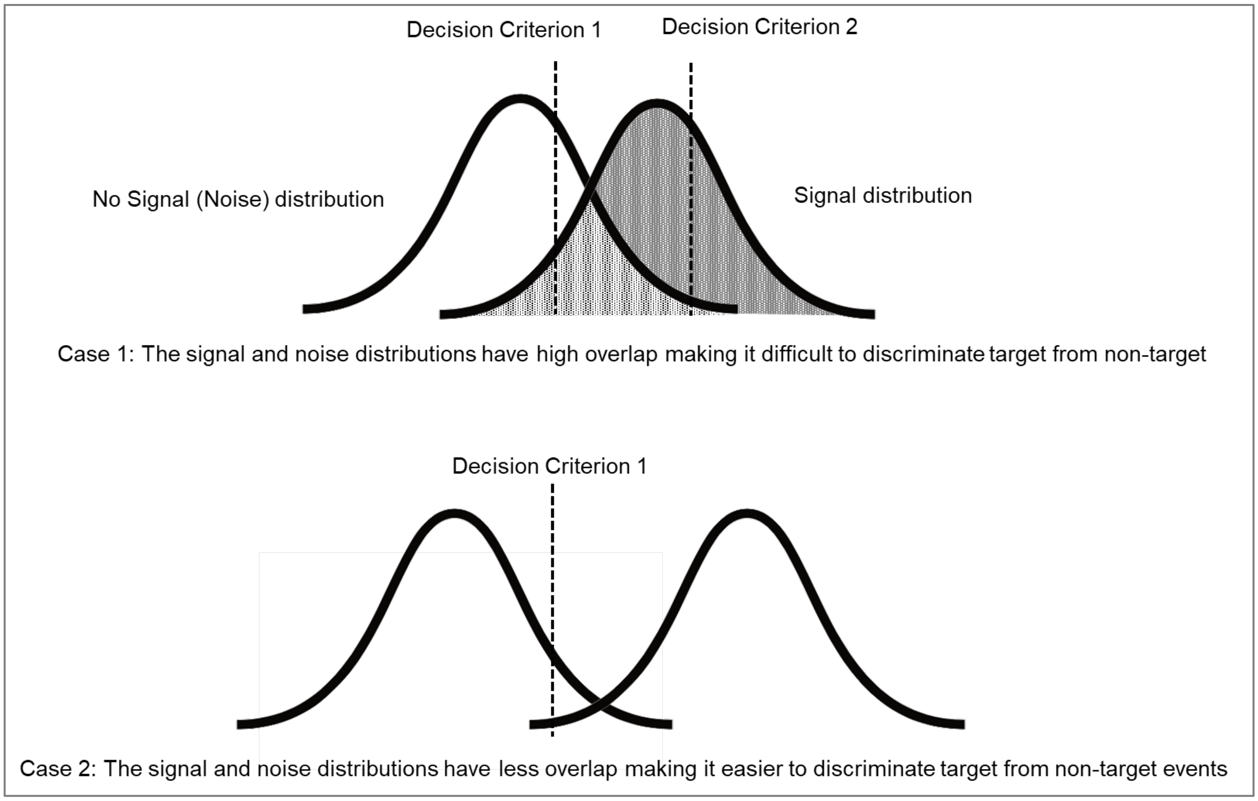

Figure 9. An illustration of the role of distribution overlap and decision criterion in Signal Detection Theory.

SDT divides the decision task into two phases: (1) an information aggregation phase where evidence relevant to the decision is gathered and integrated and (2) a decision phase where the aggregated evidence is compared to a decision criterion to determine whether the evidence is sufficient to decide that the signal (the target of interest) is present. The placement of the decision criterion is important because it determines both the 'hit' and 'false alarm' rate. Hits are instances where there is a signal, and it is correctly detected. False alarms are cases when there is no signal, but the decision maker indicated that there was. Figure 9 illustrates the fact that changing the placement of the Decision Criterion (from Criterion 1 to Criterion 2) can dramatically change the hit versus false alarm rates.

SDT can be used to distinguish people's ability to accurately detect and aggregate information (i.e., the sensitivity with which signals can be discriminated from noise) from their decision bias (i.e., where they decide to place the decision criterion) by conducting experiments (McCarley & Benjamin, 2013). Such studies have shown that placement of the decision criterion is affected by base rate (e.g., signals that are very rare may be missed) as well as the cost-benefit associated with different decision outcomes. For example, if the cost of a miss is high (e.g., missing a weapon in luggage) while the cost of a false alarm is considered low (e.g., requesting additional inspection for a possible weapon) then the decision maker is likely to use a more liberal criterion for deciding that a signal is present. On the other hand, if the cost of a false alarm is high (e.g., shooting down a civilian aircraft that was mistaken for a threat) then a stricter decision criterion is likely to be set.

Importantly, both the ability to discriminate signal from noise and the setting of the decision criteria will vary with level of expertise. For example, Parasuraman (1986) showed that experienced radiologists were better at discriminating normal from abnormal chest x-rays than radiology residents. Experienced radiologists also used a more optimal decision criterion.

Implications for Evaluation Methods and Measures
Signal detection theory provides an important methodology to use in evaluating new technologies that are intended to support the detection and sensemaking aspects of decision making. It is particularly relevant in evaluating new alerting systems to understand the hits, misses, false alarms, and correct rejection rates of the automated aid, and how those in turn influence the decisions of the human decision makers.

One of the most relevant applications of SDT is in evaluating the impact of new technologies on human decision making. SDT is particularly useful in evaluating systems that automatically generate alerts to notify a person of a potential problem that they then need to check out. The response criterion used by the automation to determine when to present the alarm can powerfully impact the human operator's trust in the system and overall performance (Meyer & Lee, 2013). For example, Dixon, Wickens, and McCarley (2007) showed that automation with high rates of false alarms hurt overall performance more than automation that had a high miss rate. A bias towards false alarms can result in poorer joint system performance by causing a shift in the decision criterion of the human decision maker (the 'cry wolf' effect). Thus, it is important to consider the impact on overall performance of the joint system when setting the decision criterion for automated agents.

LENS Model

The LENS model (Brunswik, 1955) is a mathematical tool that can be used to examine how individuals use cues in the environment to make judgments (e.g., to predict a thunderstorm from a doppler radar display; to identify a threat of aircraft collision based on indicators on a display; to decide whether a candidate is right for a job). In this sense it provides a formal way to model the contents of SA Levels 1 - 3 (Kirlik & Strauss, 2006). The LENS model starts with a characterization of the cues available in the environment for predicting the event of interest. If the available cues are imperfectly related to the event of interest (e.g., if characteristics of a job candidate are not good predictors of their future job performance) then it is not surprising that people will not be able to make accurate judgments. For example, the inherent uncertainty of meteorological phenomena and limitations of sensors can limit the reliability of cues available to weather forecasters. On the other hand, there may be cues that are good predictors of the environment but the person making the decision uses different, less reliable cues, or combines the cues in a suboptimal manner. The LENS model can be used to identify what cues a person is using and how they are combining those cues to make a judgment (called the judgment policy).

Figure 10 depicts the elements of the LENS model. The top panel shows that there is an environmental criterion (the true state of the world) that a human judge is attempting to assess based on the available cues in the environment. The cues may vary in reliability and the human judge may or may not appropriately select and combine cues. The lower panel indicates that there may be multiple judges (multiple people and/or technological systems that generate judgements). The different judges may differ in how they select and combine cues.

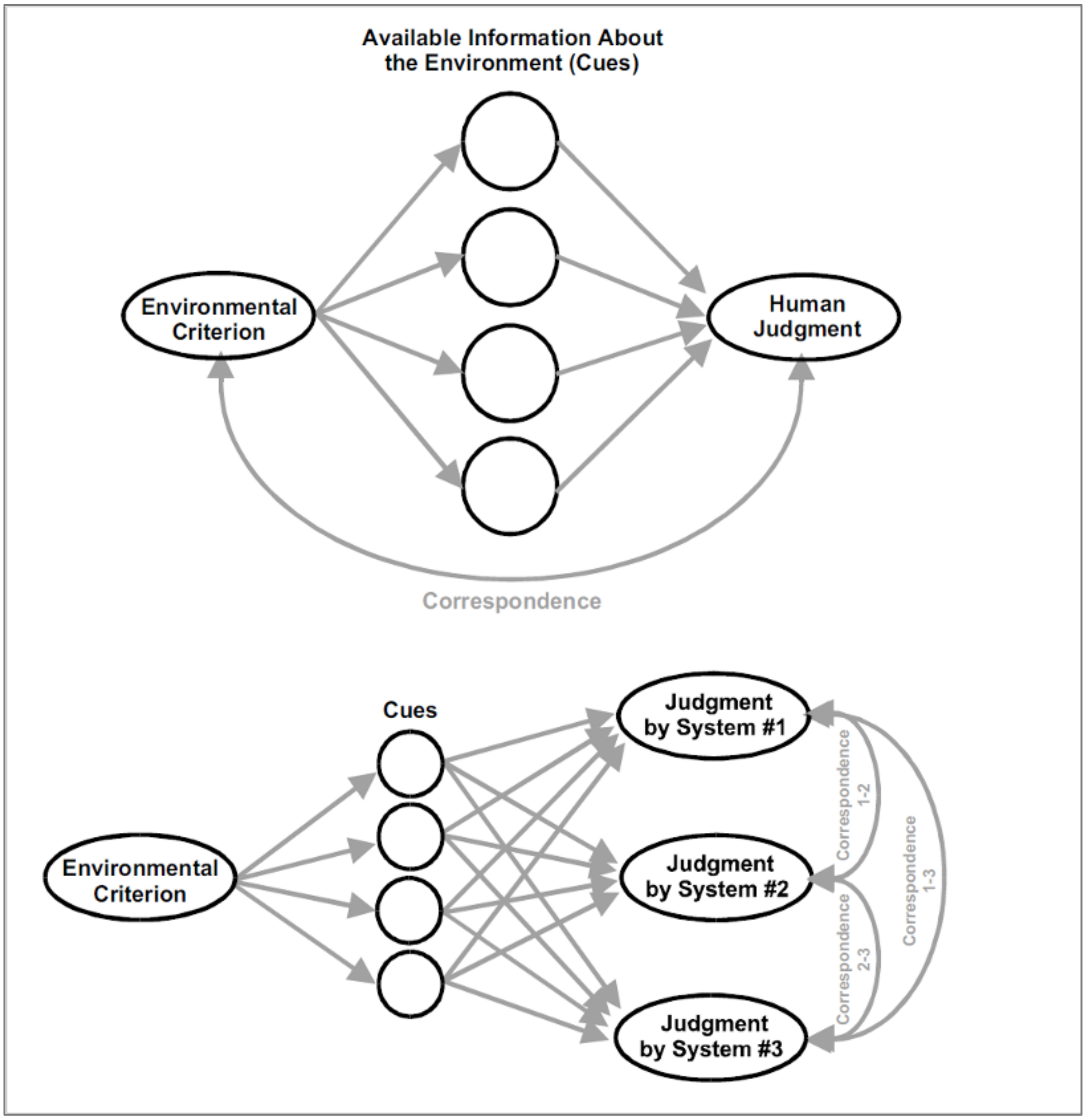

Figure 10. LENS model describing the relationship among available cues, their reliability, and how they are used in making judgments. The top panel represents a single human decision maker. The bottom panel describes the case when there are multiple (human or machine) decision makers (from Bisantz & Pritchett, 2003). Reprinted by permission of SAGE Publications.

Using the LENS model framework, multiple regressions are computed to specify (1) the predictability of the environment (i.e., the degree to which the environment can be predicted by available cues); (2) the accuracy of the decision maker's judgment policy (i.e., the extent to which the decision maker is weighting and combining cue values in a way that is similar to the "true" relationships); and (3) the judgment consistency (i.e., how consistent the decision maker is in applying the judgment policy) as well as the inter-relationship between them.

The LENS framework has been used to model human decision making, compare the judgments of automated systems with the judgements of human decision makers, examine the impact of the output of an automated system on the judgments of human decision makers, as well as examine decision making in teams (See Mosier & Fischer, 2010 for a review). For example, Bisantz and Pritchett (2003) used the LENS model to compare participant judgments of aircraft conflict detection based on cues available on a cockpit display to judgments made by two different automated conflict detection algorithms. Participant judgments were most similar to those made by a simple algorithm and most different from a more sophisticated (and generally more accurate) algorithm. The fact that the more sophisticated algorithm produced different judgments than people did, brings up the question of how to ensure that people will trust and accept the automated algorithms' judgments. The authors concluded that in order for people to trust and use the more sophisticated aid, they would need training, better displays, and/or on-demand explanation of the machine agent's decision.

Implications for Evaluation Methods and Measures

The LENS model provides a mathematical tool that can be used to design and analyze the results of experiments to answer questions such as: what cues are people using, how are they combining the cues to come up with a judgment, and are they reliable in their application of those 'rules'?

It can be a useful complement to more qualitative methods such as CTA that have been used to identify the cues that are used by experts versus less experienced individuals. In a similar way, it provides a formal method to analyze the contents of SA Levels 1 and 2 in that it identifies what cues a decision maker is picking up and how they are combining them to draw a conclusion.

The LENS model can also be used to examine the impact of new displays and decision aids on what cues people use and how that impacts their decisions.

Pulling it All Together

Looking across the set of models of decision making reviewed above, we see many areas of overlap and commonality. In this section we start by summarizing some of the core concepts that arise across models that reflect shared beliefs across the Human Factors, Naturalistic Decision Making, and Cognitive Engineering communities about decision making in dynamic environments and what is important to include in a model of decision making. We also examine implications across models for evaluation methods and measures. We end by presenting a synthesized model of decision making that we believe will provide a useful framework for characterizing decision making in dynamic environments, selecting and adapting methods and measures for evaluating decision-making in dynamic environments, and developing strategies for assessing the impact of new technologies on decision making.

Core Concepts Across Models of Decision-Making

While the various models we reviewed sometimes use different terminology to refer to similar ideas, there are clear commonalities across the models in terms of core concepts. The core concepts identified across models include:

- Decision making can arise from intuitive processes, deliberative processes, or a combination of both.
- Expert performance is often based on more intuitive, recognition-primed processes.
- Decision making is a dynamic, cyclic process, inextricable from other cognitive activities that feed into it and that it in turn influences (e.g., perception, sensemaking, planning).
- Perception involves both top-down (i.e., searching for information based on expectations) and bottom-up processes (i.e., detecting salient information in the environment that then influences comprehension and feeds further expectations).
- People actively try to understand the current situation (i.e., sensemaking) and that understanding is central to decision making.
- People's emergent understanding of a situation is guided by their knowledge and experiences (e.g., recall of past experiences, frames, mental models, schemas, scripts).
- Sensemaking often involves active construction of a mental representation of the current, unique, situation sometimes called a situation model.
- People generate expectations about what is likely to occur in the future based on their understanding of the current situation.
- People develop, revise, and adapt plans as their understanding of the current situation evolves.
- Effective teamwork requires a shared understanding of the current situation and goals, sometimes called shared SA or common ground.

The set of core concepts are summarized in Table 1. In each case we indicate which of the models of decision making include that concept, and what term the model uses for that concept as applicable. Blank cells indicate that the model does not address this idea. Our proposed framework for characterizing decision making in the context of Army Aviation builds on these core concepts. The last column in Table 1 indicates the term we will use for that concept in our proposed framework.

It will be noted that we did not include the Signal Detection Theory and the LENS model in Table 1 because they are highly specialized mathematical models with a narrow focus of coverage. While they are not included in Table 1, they remain important because they provide measurement tools to assess the results of decision-making processes, and most particularly, how new technologies and decision-support tools impact decision making. We include them in the next section where we discuss core contributions of decision-making models for evaluating decision making.

Table 1. Core concepts across different models of decision making

Core Concept	Two-System Model	RPD	Model of Situation Awareness	OODA	Decision Ladder Model	Macrocognition	Synthesized Framework
Intuitive and Deliberative Decision-Making	✓ System 1 and System 2	✓	✓	✓	✓	✓	✓
Expert decision-making often based on more automatic, intuitive judgments	✓	✓ RPD			✓	✓	✓
Decision-making as a dynamic cyclic process			✓	✓	✓	✓	✓
Perception as a top-down and bottom-up process (e.g., expectations direct attention)			✓ (Level 1 SA)	✓	✓	✓ Sensemaking	✓ Top-down and Bottom-up Processing
Situation understanding is central to decision-making		✓	✓ Comprehension (Level 2 SA)	✓ Orient	✓	✓ Sensemaking	✓ Sensemaking
Long term memory representations influence situation understanding	✓	✓	✓ Mental Models, Schemata		✓ Mental Model	✓ Mental Model, Frame, Schemata	✓ Mental Model
The output of situation understanding is a mental representation of the current situation			✓ Situation Model	✓ Mental Image	✓ Dynamic World Model	✓ Mental Model, Situation Model	✓ Situated Model
A person uses their current understanding to make projections about the future		✓ Mental Simulation	✓ Projection (Level 3 SA)	✓	✓	✓ Mental Simulation	✓ Mental Simulation
Plans are revised and adapted as understanding of the situation evolves				✓	✓ Knowledge-based Planning	✓ Adaptation and Replanning	✓ Adaptation and Replanning
Team performance relying on a shared understanding of the situation			✓ Shared SA			✓ Common Ground	✓ Common Ground

Core Contributions for Evaluating Decision Making

The various decision-making models we have reviewed contribute important perspectives, methods, and measures for evaluating decision making, as well as for evaluating the impact of new technologies on decision making for both individuals and teams. Most particularly, all of the decision-making models we reviewed emphasize the importance of examining decision making under realistic conditions that reflect the challenges that arise in the real-world situations of interest. The models have also stimulated research and contributed methods and measures for examining the impact of new technologies on decision making.

Many of the models have made methodological contributions to the design and conduct of studies evaluating decision making and assessing the impact of new technologies on decision making. Most particularly they have emphasized the need to create study conditions (e.g., through design of evaluation scenarios) that allow important aspects of decision making to be observed and measured. Many of the models have also stimulated new measures to use in evaluating decision making. The best documented and most widely used set of new measures have come out of the literature on SA, but other models of decision making have led to additional new measures as well.

We summarized the primary contribution of each decision-making model to methods and measures in the section describing that model. Here we synthesize across the models to highlight the variety of methods and measures that come out of the different decision-making modeling traditions. We anticipate that all these methods and measures will be relevant to the design of evaluation studies of decision making in the Army FVL context, and most particularly to the evaluation of the impact of new technologies on decision making in the FVL context.

In Table 2 we list some of the key contributions to study methodology/study design that have been made, and the decision-making models that are most closely associated with these methodological considerations.

In Table 3 we list some of the key contributions to evaluation measures, and the decision-making models that are most closely associated with use of these measures.

In our next report we will delve more deeply into methods and measures for evaluating decision making, with particular focus on evaluating the impact of new technologies on decision making.

Table 2. Contributions of models of decision making to methods for evaluating decision making

Contribution to Methods / Study Design	**Two-System Model**	**RPD Model**	**SA Model**	**OODA Loop**	**Decision Ladder**	**Macro cognition**	**Signal Detection Theory**	**LENS Model**
Studying decision making in naturalistic environments		✓		✓	✓	✓		
Designing realistic evaluation scenarios representative of the situations of interest	✓	✓	✓	✓	✓	✓	✓	✓
Comparing decision making of experts and novices		✓				✓	✓	✓
Comparing decision making under different technology conditions			✓		✓	✓	✓	✓
Consideration of situations requiring different modes of decision making (e.g., RPD vs. Deliberative)	✓				✓	✓		
Consideration of decision making in situations that challenge sensemaking					✓	✓		
Consideration of decision making in situations that require plan revision and adaptation					✓	✓		
Evaluating decision making processes under dynamic conditions		✓	✓	✓	✓	✓		
Using situational probes (e.g., an unexpected alarm or perceptual cue) to assess different aspects of decision making						✓		

Table 3. Contributions of models of decision making to measures for evaluating decision making

Contribution to Measures	Two-System Model	RPD Model	SA Model	OODA Loop	Decision Ladder	Macro cognition	Signal Detection Theory	LENS Model
Measures of attention allocation						✓		
Measures to identify cues used by decision makers and how they are combined to form judgements			✓					✓
Measures to differentiate perceptual discrimination from decision criterion							✓	
Measures of SA Levels 1, 2, and 3			✓					
Measures of sensemaking (e.g., assessing a person's situation model)						✓		
Measures of mental models that inform sensemaking						✓		
Measures of course of action development and revision						✓		
Measures of ability to keep pace with dynamically changing situations						✓		
Measures of decision speed and quality				✓		✓		
Measures of human-technology interaction						✓		
Team-oriented measures, including team SA, shared SA (common ground), and quality of communication and coordination			✓			✓		

Synthesized Framework for Characterizing Decision Making

In this section we present a synthesized framework that builds on the core concepts we identified across the models of decision making. The goal is to create a framework that can be used to characterize different decision-making situations that can arise in the FVL environment, assess how new proposed technologies are likely to impact different aspects of decision making, and identify appropriate methods and metrics that can be used to evaluate how new technologies affect decision making.

Our proposed framework is depicted in Figure 11. It draws most heavily on the concepts and terminology of macrocognition because this perspective provides the broadest coverage of concepts relevant to decision making in dynamic environments such as military decision making (See Table 1) as well as broadest coverage of methods and measures appropriate to evaluate decision making in dynamic environments such as FVL (see Tables 2 and 3).

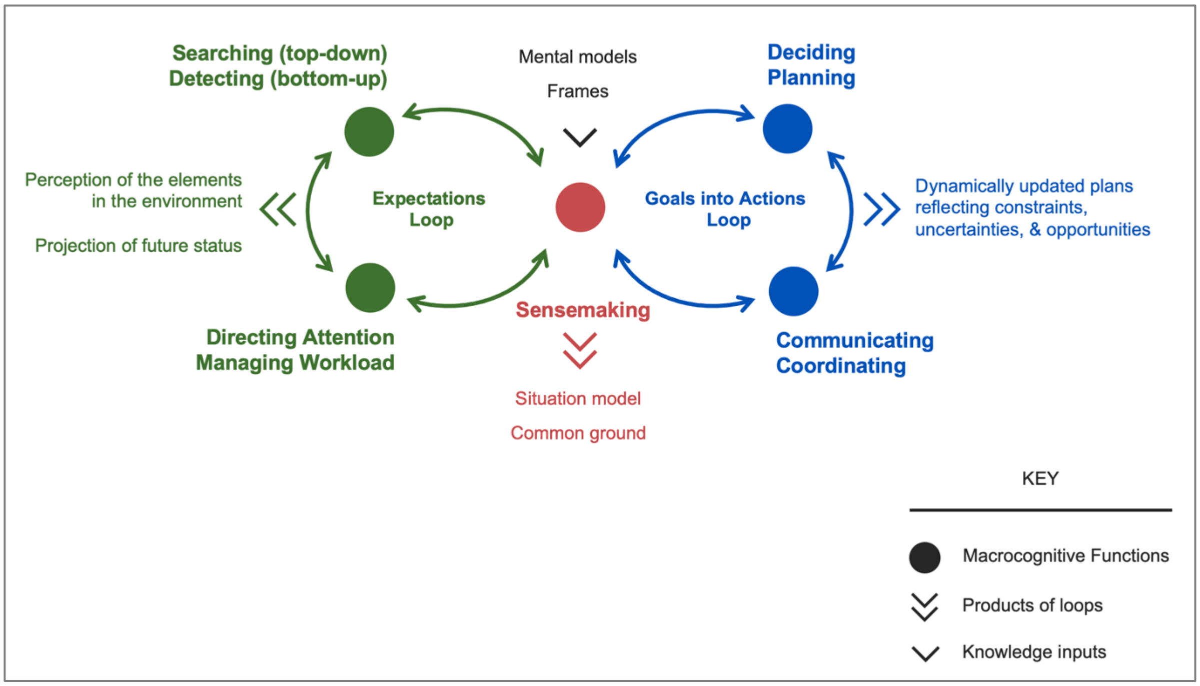

Figure 11. A synthesized framework for characterizing decision making.

We present the macrocognitive functions most relevant to dynamic decision making. In the context of this report, we do not attempt to provide a detailed process model of how these functions are accomplished. Consistent with the macrocognitive perspective, we view these functions as relevant to decision making of both individuals and teams. Most particularly we have included macrocognitive functions specifically relevant to effective team functions (e.g., communication, coordination, and the need for shared sensemaking to maintain common ground).

The macrocognitive functions we have chosen to highlight are presented in Figure 11 (represented as shaded circles in the figure). They include functions related to perception (Search & Detecting), functions related to managing attention and workload (Directing attention, Managing workload), the sensemaking function that serves as the core, functions related to formulating action plans (Deciding, Planning) and functions related to teamwork (Coordinating, Communicating).

We have chosen to highlight these functions because they are particularly relevant to decision making in complex, high-risk dynamic environments (Patterson & Hoffman, 2012; Roth et al., 2012). We acknowledge that there are multiple additional macrocognitive functions that have been documented in the macrocognitive literature that may also be relevant depending on the context but believe that this core set of macrocognitive functions will be most useful in characterizing decision making, assessing the impact of new technologies on different aspects of decision making, and identifying evaluation methods and measures for the Army's FVL environment.

The macrocognitive functions are linked together via two loops centered around sensemaking. We use the phrase 'loops' to indicate the dynamic, continuous cycle aspect of decision making, with the outputs of different macrocognitive functions continuously feeding each other. For example, sensemaking may lead to active search for information, the results of which may cause a revision in understanding of the situation, which in turn may lead to a change in goals and adaptive replanning.

Sensemaking (discussed extensively above) is the act of taking new data and figuring out how it fits together into a coherent understanding. Sensemaking may occur in a recognition-primed manner (i.e., the individual recognizes the situation based on prior experience) or it may require deliberative processes to construct an understanding drawing on long-term knowledge (e.g., frames and mental models) and the information available in the environment. Sensemaking is guided by knowledge in long-term memory (e.g., mental models, frames, schemata).

We believe that sensemaking is fundamental to decision making in the complex dynamic world of military aviation, where situations change rapidly, and pilots must constantly work to understand the current situation so as to decide on and prioritize actions. In other problem spaces, sensemaking may not have as central a role, for example in an environment where perceptual discrimination is the most critical part of the task. But for FVL, when making a tough decision, pilots are engaged in many macrocognitive functions at a time and sensemaking is at the core of many of them. Supporting sensemaking and the multiple functions around sensemaking can help yield better decisions.

An output (product) of sensemaking is an understanding of the current situation which we refer to as a ***situation model***. This is similar to the concept of comprehension (Level 2 SA) in the Endsley SA model. In cases when teams of people (and/or autonomous machine agents) are involved, there is also a need to actively build and maintain ***common ground*** across the individuals and automated agents that make up the team. Common ground depends on effective communication across team members and can be thought of as another product of sensemaking.

Sensemaking provides the basis for expectations that in turn drive attention management, workload, and search for information (***the expectations loop***). Sensemaking also provides the basis for identifying goals that in turn drive decisions and planning activities (***the turning goals into actions loop***). In the case of teams, sensemaking also influences communication and coordination across team members.

The expectations loop relates to figuring out what is important and what it means. It includes two inter-related foci:

- Searching (top-down) / Detecting (bottom-up)
- Directing attention / Managing workload

A key part of macrocognition is noticing when something is important that requires attention, something that is different and noteworthy. This can happen in two ways: when a person is intentionally monitoring or searching for data (also known as ***top-down processing***), or when a person unexpectedly detects or notices something out of the ordinary, often based on curiosity (or ***bottom-up processing***). In Boyd's (1976) OODA Loop, this was the Observe part of the loop, and he posited that the higher the quality of data gathered here, the better the decisions.

Top-down search occurs when individuals create expectations about what is going on based on their situation model and then search for data to confirm or refute those expectations. For example, a pilot might actively search for an enemy aircraft that may be hidden based on observation of other enemy aircraft and knowledge of enemy doctrine. Individuals will also make projections of the future based on their current situation model (Level 3 SA). These projections will also drive top-down perceptual processes. For example, a pilot might project when an aircraft might collide with an observed obstacle given the current situation and projected movements in the future. This will influence where they look as well as what compensatory actions they decide to take.

Bottom-up processing occurs when data in the environment draws attention. The new data may be consistent with the individual's understanding of the situation, or it might be surprising (anomalous). If the data is judged to be anomalous it will pique curiosity and trigger sensemaking processes to understand it. This in turn will lead to revising the situation model.

A related aspect is deciding how to allocate attentional resources (***directing attention***) and ***managing workload***. Determining how to allocate attention and mental workload is a critical part of macrocognition. Neither attention nor mental workload (also known as cognitive workload) are unlimited, and a key part of decision making is how attention is directed. If we are paying attention to one thing, we may not notice other important events as in the famous gorilla study where individuals failed to notice a gorilla walking across a basketball court because they were focused on counting the number of passes that the team wearing white made (Simons & Chabris, 1999). How to prioritize attention and workload is strongly influenced by an individual's current understanding of a situation and priorities.

Products of top-down and bottom-up perceptual processes include ***perception of elements in the environment*** (Level 1 SA) as well as expectations based on ***projections into the future*** (Level 3 SA).

The ***turning goals into actions loop*** relates to converting goals generated based on the current understanding of the situation into plans and actions. It includes two inter-related foci:

- Deciding / Planning
- Communicating /Coordinating

Based on a person's understanding of a situation, individuals (and teams) will define and prioritize goals, which are then converted to decisions, plans, and action execution. ***Deciding*** covers situations where decisions happen rapidly without need for deliberative planning activity. The RPD work argues that experts are able to act quickly and intuitively and make great decisions, given their wealth of experience (Klein, 1993). ***Planning*** refers to creating, revising, and executing a plan.

In dynamic environments, planning is rarely a once and for all activity. Rather, people are constantly assessing whether their current plans are appropriate to their evolving understanding of the situation and goals and revising the plan accordingly. In many cases this requires *adapting the plan* or even coming up with an entirely new plan on the fly. Also critical in a military aviation environment is the ability to recognize opportunities for action that had not been anticipated ahead of time and exploiting them. This is sometimes referred to as *turning leverage points into courses of action* (Klein et al., 2003).

As Patterson and colleagues (Patterson, Miller, Roth, & Woods, 2010) discussed, planning can include: "Adaptively responding to changes in objectives from supervisors and peers, obstacles, opportunities, events, or changes in predicted future trajectories; when ready-to-hand default plans are not applicable to the situation, this can include creating a new strategy for achieving one or more goals or desired end states." (pg. 272). Often, an initial plan must be tweaked as more information comes in or as the decision maker assesses success. Replanning builds on the initial plan and can be thought of as a continuous cycle.

Another important element of planning is *managing uncertainty*. Various options for managing uncertainty include collecting more information to reduce uncertainty and/or deciding on actions that will accommodate a range of possibilities (e.g., mitigating the worst possible outcome). Also important is *managing risks and trade-offs*. In our own work we have seen Army aviators decide to incur more risk for themselves in order to reduce overall risk to the mission (e.g., attempting a riskier landing in order to reduce risk to forces once they are on the ground or to better support actions on the objective).

Note that the output of planning can itself lead to revision in understanding the situation, as well as top-down search for more information highlighting the cyclic nature of the goals loop.

The second focus of the goals to actions loop covers teamwork functions. This includes ***communicating,*** which is critical to maintaining common ground, aligning goals, and synchronizing activity. Also included is ***coordinating*** - determining how to coordinate action across team members to achieve a shared goal.

Products of the goals to action loop include dynamically updated plans that reflect constraints, uncertainties, and opportunities.

Conclusions

In this review of the literature on decision making, we surveyed models from many different academic and applied philosophies to synthesize an approach that we hope will be helpful as we consider methods and measures for evaluating decision-support tools. Highlighting the functions needed to support good decision making will inform the selection of methods and measures that evaluate those functions in particular. For example, the Situation Awareness researchers (Endsley, 1995b) have developed methods focused on assessing aspects of the Expectations loop—perceptions of the elements in the environment, and how people project their future status—as well as facets of sensemaking. Researchers in macrocognition have developed methods and measures of macrocognitive functions (see Patterson & Miller, 2010, for many examples).

In the *Core Contributions for Evaluating Decision Making* section of this report, we identified literatures that offer practical methods and measures for assessing the elements articulated in Figure 11, *Synthesized framework for characterizing decision making*. Our next report will focus on detailing these methods and measures, with an emphasis on tailoring them for use in assessing the impact of proposed technologies in the context of Future Vertical Lift.

Our goal moving forward is to explore both methods for evaluation, to ensure that the process represents the context of intended use, as well as specific measures. From our literature review of decision-making models, it is clear that we must explore how we structure the evaluation (study method) as well as what we measure. We will explore these and other approaches to measurement and methodology in order to elucidate how to create scenarios for evaluation and what to look for during the evaluation process.

References

Angerman, W. S. (2004). *Coming Full Circle with Boyd's OODA Loop Ideas: An Analysis of Innovation Diffusion and Evolution*. Theses and Dissertations. 4085.

Beach, L. R. (1993). Broadening the definition of decision making: The role of prochoice screening of options. *Psychological Science*, 4 (4) 2015 – 2020.

Bisantz, A. and Roth, E. M. (2015). Decision Making and Human Systems Integration. In Boem-Davis, D. A., Durso, F. T and Lee, J. D. (Eds.) *APA Handbook of Human Systems Integration*. Washington, DC: American Psychological Association.

Bisantz, A. M., & Pritchett, A. R. (2003). Measuring the fit between human judgments and automated alerting algorithms: A study of collision detection. *Human Factors*, *45*(2), 266-280.

Boyd, J. (1987). *A discourse on winning and losing*. Maxwell Air Force Base, AL: Air University Library Document No. M-U 43947 (Briefing slides)

Boyd, J. R. (1976). Destruction and Creation. Unpublished manuscript. Cited in Angerman, William S., *Coming Full Circle with Boyd's OODA Loop Ideas: An Analysis of Innovation Diffusion and Evolution*. (2004). Theses and Dissertations. 4085.

Brehmer, B. (2005). The dynamic OODA loop: Amalgamating Boyd's OODA loop and the dynamic decision loop.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 62, 193 - 217.

Byrne, E. (2015). Commentary on Endsley's "Situation Awareness Misconceptions and Misunderstandings." *Journal of Cognitive Engineering and Decision Making*. 9(1), 84-86. doi:10.1177/1555343414554703

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

Chang, Y.J., Bley, D., Criscione, L., Kirwan, B. Mosleh, A., Madary, T., Nowell, R., Richards, R., Roth, E., Sieben, S. & Zoulis, A. (2014). The SACADA database for human reliability and human performance. *Reliability Engineering and System Safety*. 125, 117-133.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). American Psychological Association. https://doi.org/10.1037/10096-006

Crandall, B., Klein, G. & Hoffman, R. (2006). *Working Minds: A Practitioner's Guide to Cognitive Task Analysis*. Bradford Publishing Company, Bradford, PA.

Dixon, S. R., Wickens, C. D. & McCarley, J. S. (2007). On the independence of compliance and reliance: Are automation false alarms worse than misses? *Human Factors* 49(4), 564-572.

Endsley M. R. (1995b). Measurement of Situation Awareness in Dynamic Systems. *Human Factors*. 37(1):65-84.

Endsley, M. R. (2015). Situation Awareness Misconceptions and Misunderstandings. *Journal of Cognitive Engineering and Decision Making*. 9(1), 4-32. doi:10.1177/1555343415572631

Endsley, M. R. (2000). Theoretical underpinnings of situation awareness: A critical review. In Endsley, M. R & Garland, D. J. (Eds). *Situation Awareness Analysis and Measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.

Endsley, M. R. (2021a). Situation Awareness. In Salvendy G and Karwowski, W. (Eds). *Handbook of Human Factors and Ergonomics*. NJ: John Wiley & Sons.
Endsley, M. R. (2021b). A systematic review and meta-analysis of direct objective measures of situation awareness: a comparison of SAGAT and SPAM. *Human Factors*, *63*(1), 124-150.
Endsley, M. R. & Jones, W. M. (2001). A model of inter and intra team situation awareness: Implications for design, training and measurement. In McNeese, M., Salsa, E. & Endsley, M. (Eds). *New Trends in Cooperative Activities: Understanding System Dynamics in Complex Environments*. Santa Monica, CA: Human Factors and Ergonomics Society.
Endsley, M.R. (1995a). Toward a theory of situation awareness in dynamic systems. *Human Factors*. 37 (1): 32–64.
Ernst, K., Militello, L., & Scheff, S. (2021). *Future Vertical Lift Complexity in Information Systems Conceptual Dynamic World Data Model report* [unpublished technical report]. Applied Decision Science.
Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment and social cognition. *Annual Review of Psychology*, 59, 255-278.
Fadok, D. S. (1995). John Boyd and John Warden -- Air Power's Quest for Strategic Paralysis. Air University, Maxwell AFB, AL
Fischhoff, B. & Broomell, S. (2020). Judgment and Decision Making. *Annual Review of Psychology*, 71, 331-55.
Gigerenzer, G. & Gaissmaier, W. (2011). Heuristic Decision Making. *Annual Review of Psychology*, 62, 451-482.
Gigerenzer, G., Todd, P. M., & the ABC Research Group, (Eds.). (1999). *Simple heuristics that make us smart*. Oxford: Oxford University Press.
Gonzalez, C. & Meyer, J. (2016). Integrating trends in decision-making research. *Journal of Cognitive Engineering and Decision Making*, 10 (2), 120-122.
Green, D. M. & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Oxford: Wiley
Hoffman, R. (2010). Some Challenges for Macrocognitive Measurement. In Patterson E.S., & Miller J. (Eds.) *Macrocognition Metrics and Scenarios: Design and Evaluation for Real-World Teams*. Ashgate Publishing.
Hoffman, R. & Yates, J.F. (2005). Decision Making. *IEEE Intelligent Systems*. 20:22-29.
Kahneman, D. (2011). *Thinking fast and slow*. New York, NY: Farrar, Straus, & Giroux.
Kahneman, D. & Klein, G. (2009). Conditions for intuitive expertise. *American Psychologist*, 64 (6), 515 – 526.
Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47 (2), 263-292.
Kirlik, A., & Strauss, R. (2006). Situation awareness as judgment I: Statistical modeling and quantitative measurement. *International Journal of Industrial Ergonomics*, *36*(5), 463-474.
Klein, G. (2015). Whose Fallacies? *Journal of Cognitive Engineering and Decision Making*. 9(1), 55-58. doi:10.1177/1555343414551827

Klein, D.E., Klein, H.A. and Klein, G. (2000). "Macrocognition: Linking Cognitive Psychology and Cognitive Ergonomics," *Proc. 5th Int'l Conf. Human Interactions with Complex Systems*, Univ. of Illinois at Urbana-Champaign , pp. 173–177.

Klein, G. (1989). Recognition-primed decisions. In W. B. Rouse (Ed.) *Advances in Man-Machine Systems Research*, vol. 5, Greenwich, CT:JAI Press, 47-92.

Klein, G. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G. Klein, J. Orasanu, R. Calderwood and C. E. Zsambok (Eds). *Decision Making in Action: Models and Methods*. Norwood, NJ: Ablex Publishing.

Klein, G. (2007a). Flexecution as a paradigm for replanning, Part 1. *IEEE Intelligent Systems*, 22(5), 79-83.

Klein, G. (2007b). Flexecution, Part 2: Understanding and supporting flexible execution. *IEEE Intelligent Systems*, 22(6), 108-112.

Klein, G. (2010). Macrocognitive Measures for Evaluating Cognitive Work. In Patterson E.S., & Miller J. (Eds.) *Macrocognition Metrics and Scenarios: Design and Evaluation for Real-World Teams*. Ashgate Publishing.

Klein, G. (2011). *Streetlights and Shadows: Searching for the keys to Adaptive Decision Making*. The MIT Press: Cambridge, MA.

Klein, G. & Wright, C. (2016). Macrocognition: From Theory to Toolbox. *Frontiers in Psychology* (7). 54.

Klein, G., Feltovich, P., Bradshaw, J. M., Woods, D. D. (2005). Common ground and coordination in joint activity. In Rouse, W., Boff, K. (Ed.). *Organizational Simulation*, Wiley, pp. 139–178.

Klein, G., Moon, B., & Hoffman, R. R. (2006). Making sense of sensemaking 1: Alternative perspectives. *IEEE Intelligent Systems*, 21, 7073.

Klein, G., Phillips, J., Rall, E., & Peluso, D. (2007). A Data-Frame theory of sensemaking. *Expertise out of Context: Proceedings of the Sixth International Conference on Naturalistic Decision Making* (p. 113-155). Psychology Press.

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

Klein, G.A., Wolf, S., Militello L., & Zsambok, C. (1995). Characteristics of skilled option generation in chess. *Organizational Behavior and Human Decision Processes*, 62(1), 63-69

Levin, I. P., Schneider, S. L. & Gaeth, G. J. (1998). All frames are not equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes*, 76 (2), 149 – 188

Lintern G. A. (2010). Comparison of the Decision Ladder and the Recognition-Primed Decision Model. *Journal of Cognitive Engineering and Decision Making*. 4(4), 304-327

Lipshitz, R. (1993). Converging themes in the study of decision making in realistic settings. *Decision making in action: Models and methods*, 103-137.

McCarley, J. S. and Benjamin, A. S. (2013). Bayesian and signal detection models. In Lee, J. D. & Kirlik, A. (Eds) *Oxford Handbook of Cognitive Engineering*. New York, NY: Oxford University Press, pp 465-475.

Militello, L. & Anders, S. (2019). Incident-based methods for studying expertise. In P. Ward, JM Schraagen, J. Gore and E. M. Roth (Eds) *The Oxford Handbook of Expertise*. UK: Oxford University Press.

Militello, L. G., Klein, G. (2013). Decision-centered design. In Lee, J. D., Kirlik, A. (Eds.), *The Oxford handbook of cognitive engineering*. Oxford, UK: Oxford University Press.

Militello, L. Roth, E., Scheff, S., Ernst, K., Sushereba, C., Klein, D., & Wonderly S. (2019). *Crew Configuration Analysis for Future Airborne Reconnaissance Operations*. Applied Decision Science.

Militello, L., Roth, E., Scheff, S., Ernst, K., Sushereba, C., Diiulio, J. and Klein, D. (2018). *Toward an optimally crewed future vertical lift vehicle: Overview of the envisioned world, core missions, and pertinent technology*. Applied Decision Science

Morrison, J. G., Kelly, R. T., Moore, R. A, & Hutchins, S. G. (1998). Implications of decision-making research for decision support and displays. In J. A. Cannon-Bowers & E. Salas (Eds.), *Making Decisions under Stress* (pp. 375 - 406). Washington, DC: APA.

Mosier KL, Fischer UM. (2010). Judgment and Decision Making by Individuals and Teams: Issues, Models, and Applications. *Reviews of Human Factors and Ergonomics*. 6(1):198-256.

Mumaw, R., Roth, E.M., Vicente K., & Burns, C. (2000). There is more to monitoring a nuclear power plant than meets the eye. *Human Factors*. Spring 2000;42(1):36-55

Naikar, N. (2010). *A comparison of the decision ladder template and the recognition-primed decision model*. Defence Science and Technology Organization, Victoria (Australia) Air Operations Division.

Oskarsson, A.T. & Hodgins, C.R. (2010). The AlphaAct decision support system for emergency responders. *Proceedings of the 54th Annual Meeting of the Human Factors and Ergonomics Society*.

Parasuraman R, Sheridan TB, Wickens CD. (2008). Situation Awareness, Mental Workload, and Trust in Automation: Viable, Empirically Supported Cognitive Engineering Constructs. *Journal of Cognitive Engineering and Decision Making*. 2(2), 140-160. doi:10.1518/155534308X284417

Parasuraman, R. (1986). Effects of practice on detection of abnormalities in chest x-rays. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. pp. 309-311. Santa Monica, CA: HFES

Patterson E.S., & Miller J. (Eds.) (2010). *Macrocognition Metrics and Scenarios: Design and Evaluation for Real-World Teams.* Ashgate Publishing.

Patterson, E. & Hoffman, R. (2012). Visualization framework of macrocognition functions. Cognition Technology and Work, 14(3) 221-227.

Patterson, E. S., Roth, E. M., Woods, D. D. (2010). Facets of complexity in situated work. In Patterson E.S., & Miller J. (Eds.) *Macrocognition Metrics and Scenarios: Design and Evaluation for Real-World Teams*. Ashgate Publishing. (pp. 221- 251).

Rasmussen, J. (1986). *Information processing and human machine interaction: An approach to cognitive engineering*. New York: North-Holland.

Rasmussen, J. L. (1976). Outlines of a hybrid model of the process plant operator. In Sheridan, TB and Johannsen, G, Monitoring Behavior and Supervisory Control. In *International*

Symposium on Monitoring Behavior and Supervisory Control (pp. 371-384). Plenum Publishing Corp.
Rasmussen, J., Pejtersen, A. M., & Goodstein, L. P. (1994). *Cognitive Systems Engineering*. New York: Wiley and Sons.
Roth, E. M., Sushereba, C. & Diiulio, J. (2018). *From MABA-MABA to Interdependent design: A critical review of alternative approaches to function allocation to support optimal crewing*. Applied Decision Science.
Roth, E.M., Mosleh, A., Chang, Y.J., Richards, R., Bley, D., Shen, S.H., & Zoulis, A. (2012). A Model-based Framework for Characterizing Contextual Factors for HRA Applications. *PSAM11/ESREL2012 Conference*, Helsinki, Finland, June (p 25-29).
Rousseau, R. & Breton, R. (2004). The M-OODA: A model incorporating control functions and teamwork in the OODA loop. *In Proceedings of the Command and Control Research, and Technology Symposium.*
Schraagen, J.M., Klein, G. & Hoffman, R. R. (2008). The Macrocognition Framework of Naturalistic Decision Making. In JM Schraagen, L. G. Militello, T. Ormerod, & L. Raanan (Eds) *Naturalistic Decision Making and* Macrocognition (1st ed.). CRC Press. https://doi.org/10.1201/9781315597584
Shanteau, J. (1992). Competence in experts: The role of task characteristics. *Organizational behavior and human decision processes*, 53, 252-262.
Simons, D. & Chabris, C. (1999). Gorillas in our midst: sustained inattentional blindness for dynamic events. *Perception*, 28(9), 1059-1074.
Stanton NA, Salmon PM, Walker GH. (2015). Let the Reader Decide: A Paradigm Shift for Situation Awareness in Sociotechnical Systems. *Journal of Cognitive Engineering and Decision Making*. 9(1), 44-50. doi:10.1177/1555343414552297
Tenney YJ, Pew RW. (2006). Situation Awareness Catches On: What? So What? Now What? *Reviews of Human Factors and Ergonomics*. 2(1), 1-34
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124
U.S. Department of Transportation, Federal Aviation Administration. (2014). *Facility operation and administration (Order JO 7210.3Y).* Retrieved from http://www.faa.gov/document Library/media/Order/FAC.pdf
Van Dijk, T. A. & Kintsch, W. (1983). *Strategies of Discourse Comprehension*. New York: Academic Press.
Vicente, K., Mumaw, R. & Roth, E. (2004). Operator monitoring in a complex dynamic work environment: a qualitative cognitive model based on field observations. *Theoretical Issues in Ergonomics Science*, 5(5), 359-384.
Wiggins, S. L. & Cox, D.A. (2010). System Evaluation Using the Cognitive Performance Indicators. In Patterson E.S., Miller J. (Eds.) *Macrocognition Metrics and Scenarios: Design and Evaluation for Real-World Teams*. Ashgate, UK.

U.S. Army Aeromedical Research Laboratory
Fort Rucker, Alabama

All of USAARL's science and technical information documents are available for download from the Defense Technical Information Center.

https://discover.dtic.mil/results/?q=USAARL

Army Futures Command
U.S. Army Medical Research and Development Command