Measuring Attributes of Rigor in Information Analysis
Daniel J. Zelik, Emily S. Patterson, and David D. Woods

Introduction

Information analysis describes the process of “making inferences from available data” and determining “the best explanation for uncertain, contradictory, and/or incomplete data” (Trent, Patterson, & Woods, 2007, p. 76). It is an accumulation and interpretation of evidence to support decision making (Heuer, 1999; Schum, 1987). And it represents an active, goal-directed, and often technologically mediated, sensemaking activity (Klein, Phillips, Rall, & Peluso, 2007).

This macrocognitive framing of information analysis as “sensemaking” (Klein, in press; Klein, Moon, & Hoffman, 2006a, 2006b) highlights that cognition employs narrative elements to recognize and communicate meaning, and that sensemaking—as with other macrocognitive functions such as (re)planning—is (1) joint activity distributed over time and space, (2) coordinated to meet complex, dynamic demands, (3) conducted in an uncertain, event-driven environment, (4) often performed with simultaneously complementary and conflicting goals with high consequences for failure, (5) made possible by effective expertise in roles, and (6) shaped by organizational constraints (see Patterson & Miller, in press; Klein et al., 2003; Schraagen, Klein, & Hoffman, 2008).

Thus, information analysis in both professional domains such as intelligence analysis and more everyday contexts like online shopping is a deliberate process of collecting data, reflecting upon it, and aggregating those findings into knowledge, understanding, and the potential for action. However, even in contexts where there are relatively low consequences for inaccurate conclusions, there lurk ever-present vulnerabilities exacerbated by ever increasing data availability and connectivity (Woods, Patterson, & Roth, 2002; Treverton & Gabbard, 2008; Betts, 2002, 1978). Of particular concern is the potential inherent in all analytical activity that an analysis process is prematurely concluded and is subsequently of inadequate depth relative to the demands of a situation—a vulnerability we characterize as the risk of shallow analysis.
This chapter examines rigor as a measure of sensemaking in information analysis activity that captures how well an analytical process reduces this risk. It describes eight attributes of analytical rigor that were synthesized from empirical studies with professional intelligence analysts from a variety of specialties and agencies (Patterson, Roth, & Woods, 2001; Miller, Patterson, & Woods, 2006; Trent, Patterson, & Woods, 2007; Zelik, Patterson, & Woods, 2007; Grossman, Woods, & Patterson, 2007). Examples are then provided that illustrate how these attributes were calibrated for two different studies of analytical activity. Finally, potential future directions are discussed for using the attributes to assess, negotiate, and communicate the rigor of information analysis processes.

**Analytical Rigor as a Measure of Process**

Across domains, conventional perspectives for measuring analytical rigor focus on identifying gaps between what was actually done versus a prescribed or “standard” method, with rigor variously defined as the “scrupulous adherence to established standards” (Crippen et al., 2005, p. 188), the “application of precise and exacting standards” (Military Operations Research Society, 2006, p. 4), “methodological standards for qualitative inquiry” (Morse, 2004, p. 501), and the “unspoken standard by which all research is measured” (Davies & Dodd, 2002, p. 280).

Unfortunately, such definitions suggest a conceptualization of analytical activity that is neither particularly likely to reflect rigorous analysis work as practiced (Dekker, 2005; Sandelowski, 1993, 1986), nor particularly useful in developing the concept of rigor as a measure of performance. The diverse nature of intelligence analysis in particular renders the identification of a single “standard” process an intractable, if not impossible task (Berkowitz & Goodman, 1991; Krizan, 1999; Marrin, 2007). Consequently, rather than on a standards-based notion of rigor, our measurement approach focuses on how the risk of shallow analysis is reduced via analyst-initiated strategies that are opportunistically employed throughout the analysis process. These strategies are alternatively conceptualized as “broadening” checks (Elm et al., 2005) insofar as they tend to slow the production of analytic product and make explicit the sacrifice of efficiency in pursuit of accuracy, a central tenet of the framework.
Stories of Shallow Analysis

In our research with professional information analysts, study participants were often asked to directly evaluate or critique the products and processes of other analysts (e.g., Miller, Patterson, & Woods, 2006; Zelik, Patterson, & Woods, 2007). Across these studies, consistent patterns emerged as experienced analysts identified critical vulnerabilities in information analysis by way of “stories” describing how analytical processes often can or did go wrong. We organize these stories into eight inter-related risks of shallow analysis.

Specifically, we found that shallow analysis:

1. **Is structured centrally around an inaccurate, incomplete, or otherwise weak primary hypothesis, which analysts sometimes described as favoring a “pet hypothesis” or as a “fixation” on an initial explanation for available data.**

2. **Is based on an unrepresentative sample of source material, e.g. due to a “shallow search”, or completed with a poor understanding of how the sampled information relates to the larger scope of potentially relevant data, e.g. described as a “stab in the dark”.**

3. **Relies on inaccurate source material, as a result of “poor vetting” for example, or treats information stemming from the same original source as if it stems from independent sources, labeled variously as “circular reporting”, “creeping validity”, or as the “echo chamber” effect.**

4. **Relies heavily on sources that have only a partial or, in the extreme, an intentionally deceptive stance toward an issue or recommended action, often characterized by analysts in terms of “biased”, “slanted”, “polarized”, or “politicized” source material.**

5. **Depends critically on a small number of individual pieces of often highly uncertain supporting evidence proving accurate, identified by some individuals as an analysis heavily dependent upon “hinge evidence” or, more generically, as a “house of cards” analysis.**


6. Contains portions that contradict or are otherwise incompatible with other portions, e.g. via the inclusion of lists or excerpts directly “cut and paste” from other documents or via an assessment that breaks an issue into parts without effectively re-integrating those parts.

7. Does not incorporate relevant specialized expertise, e.g. an analyst who “goes it alone”, or, in the other extreme, one who over relies on the perspectives of domain experts.

8. Contains weaknesses or logical fallacies in reasoning from data to conclusion, alternatively described as having a “thin argument”, a “poor logic chain”, or as involving “cherry picking” of evidence.

In addressing each of these sources of risk, eight corresponding attributes of analytical rigor were identified (see Figure 1). Together these attributes represent a model of how experts critique analytical process, rather than a model of the analytical activities of those experts. Consequently, these attributes make salient the aspects of process that are attended to in making an expert judgment of analytical rigor.

Our approach to operationalize the model involves translating empirically based descriptions of each attribute into differentiating categories of analytical behaviors. Each attribute is thus characterized in terms of low, moderate, and high indicators of rigorous process. Using this rubric, analytical processes can be rated or scored by comparing the methods employed during a given analysis to the categorical descriptions that appear in Figure 1. Scoring an analysis process across all attributes, in turn, reflects an overall assessment of analytical rigor.

Generally, if no broadening checks are observed in relation to a particular aspect of risk, we code it as “low” since analysts, both casual and professional, typically show some awareness of the possible risks of being inaccurate in an assessment with respect to a given attribute. For “moderate” rigor, encompassed are activities that display an active response to a particular type of risk, while strategies that are particularly effortful, systematic, or exhaustive are coded as “high.” Finally, aspects of an analysis that are not applicable to a given context are generally not scored, for example attributes involving collaboration on a task where an analyst was expected to work alone.
The focused collection of data in a critical evaluation of the analytical reasoning process.

- **Hypothesis Exploration**
  - The construction and evaluation of potential explanations for collected data.
  - Failure to go beyond routine and readily available data sources.
  - Reliance on a single source type or on data that are far removed from original sources.
  - Dependence upon "pushed" information, rather than on actively collected information.
  - Use of stale or dated source data.

- **Information Search**
  - Collection from multiple data types or reliance on proximal sources to support key findings.
  - Some active information seeking.

- **Information Validation**
  - General acceptance of information at face value, with little or no clear establishment of underlying veracity.
  - Lack of convergent evidence.
  - Poor tracking and citation of original sources of collected data.

- **Stance Analysis**
  - Little consideration of the views and motivations of source data authors.
  - Recognition of only clearly biased sources or sources that reflect a well-defined position on an issue.

- **Sensitivity Analysis**
  - Explanations are appropriate and valid at a surface level.
  - Little consideration of the critical "what if" questions, e.g., "What if a given data source turns out to be unreliable?" or "What if a key prediction does not transpire as anticipated?"

- **Information Synthesis**
  - Little insight with regard to how the analysis relates to the broader analytical context or to more long-term concerns.
  - Lack of selectivity, with the inclusion of data or figures that are disconnected from the key arguments or central issues.
  - Extensive use of lists or the restatement of material copied directly from other sources with little reinterpretation.

- **Specialist Collaboration**
  - Minimal direct collaboration with experts.
  - Little if any on-topics, "outside" expertise is accessed or sought out directly.

- **Explanation Critiquing**
  - Few if any instances of alternative or "outside" criticisms being considered.
  - Reliance on preceding channels of critiquing, primarily those supranary.

**Indicators of LOW Rigor:**
- Little or no consideration of alternatives to primary or initial hypotheses.
- Interpretation of ambiguous or conflicting data such that they are compatible with existing beliefs.
- Fixation or knowledge shielding behaviors.
- Some consideration of how data could support alternative hypotheses.
- An unbalanced focus on a probable hypothesis or a lack of commitment to any particular hypothesis.

**Indicators of MODERATE Rigor:**
- Failure to go beyond routine and readily available data sources.
- Reliance on a single source type or on data that are far removed from original sources.
- Dependence upon "pushed" information, rather than on actively collected information.
- Use of stale or dated source data.

**Indicators of HIGH Rigor:**
- Significant generation and consideration of alternative explanations via the direct evaluation of specific hypotheses.
- Incorporation of "outside" perspectives in generating hypotheses.
- Evolution and broadening of hypothesis set beyond an initial framing.
- Ongoing revision of hypotheses as new data are collected.
- Collection of data from multiple source types in addition to the use of proximal sources for all critical inferences.
- Exhaustive and detailed exploration of data in the relevant sample space.
- Active approach to information collection.

**Figure 1.** Attributes of rigorous analysis organized with respect to indicators of low, moderate, and high rigor.
Attributes of Analytical Rigor

1. **Hypothesis Exploration.** Hypotheses are among the most basic building blocks of analytical work, representing candidate explanations for available data (Elm et al., 2005) and one of the necessary components of inferential analysis (Schum, 1987). As noted by Kent (1966), “what is desired in the way of hypotheses... is quality and quantity... a large number of possible interpretations of the data, [and] a large number of inferences, or concepts, which are broadly based and productive of still other concepts” (p. 174). Consistent with this perspective, rigorous analysis is identified by the depth and breadth of the generation and consideration of alternatives, by the incorporation of diverse perspectives in brainstorming hypotheses, by the evolution of thinking beyond an initial problem framing, and by the ongoing openness to the potential for revision.

Hypothesis exploration also represents a check against “premature narrowing” (Woods & Hollnagel, 2006; Cooper, 2005), given the concern “that if analysts focus mainly on trying to confirm one hypothesis they think is probably true... [they] fail to recognize that most of this evidence is also consistent with other explanations or conclusions” (Heuer, 1999, p. 96). Accordingly, indicators of weak hypothesis exploration include a focus on confirming a single explanation, a failure to consider alternative perspectives on available data, an interpretation of ambiguous or conflicting data such that it is compatible with existing beliefs, and analyst behavior indicative of fixation (Pherson, 2005; De Keyser & Woods, 1990).

2. **Information search.** Similarly viewed as a fundamental component of analysis work, this attribute is alternatively described as the “collection of data bearing upon the [analysis] problem” (Kent, 1966, p. 157) or as the focused ”extraction of an essential, representative, 'on analysis' sample from available data” (Elm et al., 2005, p. 297). Thus, information search encompasses all activities performed to gather task-relevant evidence—including those to broaden as well as deepen, those that are active as well as passive, and those hypothesis-driven as well as data-driven. Note that this framing of information search reflects the diverse nature of analytical activity and emphasizes the fact that, for the professional analyst, supporting evidence comes in many forms, and not simply as raw data (Krizan, 1999).

Information search is primarily concerned with where and how analysts look for supporting information. A strong information search process is characterized by the extensive exploration of relevant data, by the collection of data from multiple source types, and, most critically, by an active approach
to information collection. In this context, active collection describes efforts to seek information beyond that which is readily available (Serra, 2008). A weak information search in contrast is identified by failure to go beyond routine and readily available data sources, by reliance on a single source type or on “distant” data that is removed from original source material, and by passive dependence upon “pushed” rather than actively collected data.

3. **Information Validation.** This attribute is concerned with the “critical evaluation of data” (Kent, 1966, p. 157) and with determining the level of conflict and corroboration—or agreement and disagreement—among sources (Elm et al., 2005). In rigorous analysis, analysts make an explicit effort to distinguish fact from judgment and are concerned with consistency and credibility among, as well as within, sources. Thus, a strong validation process involves assessing the reliability of sources, assessing the appropriateness of sources relative to the task question, and the use of proximate sources whenever possible (Krizan, 1999). It also involves an explicit effort to seek out multiple, independent sources of converging evidence for key findings.

In contrast, weak information validation is reflected in the uncritical acceptance of data at face value, little or no clear effort to establish underlying veracity, and a failure to collect independent supporting evidence. Poor tracking and citation of original sources also identify such analyses. Between strong and weak characterizations, a moderate validation process involves the recognition of inconsistencies among sources and, often times, involves the use of heuristics to support judgments of source integrity—such as deference to sources that have previously proven highly reliable and avoidance of those that have not. On the aggregate, then, information validation can be described as an intense concern with issues of agreement, consistency, and reliability with respect to the set of collected data.

4. **Stance analysis.** “Stance” refers to the perspective of a source on a given issue and is alternatively discussed in the language of viewpoint, outlook, or affiliation—though more often it is characterized in terms of slant, bias, or predisposition. Stance analysis refers to the evaluation of information with the goal of identifying the positions of sources with respect to a broader contextual understanding and in relation to alternative perspectives on an issue. In the extreme, this analysis serves as a guard against deception (Johnson, Grazioli, Jamal, & Berryman, 2001; Moore & Reynolds, 2006; Heuer, 1987). More prosaically however, it serves as a check
against a reliance on unrepresentative data sets that reveal only partial perspectives on an issue.

A process in which little attention is paid to issues of stance reflects weak analysis. In such instances, the analysis may identify heavily slanted sources or sources that support a well-defined position on an issue but yet reflect little in the way of a nuanced understanding. A somewhat improved stance analysis would incorporate basic strategies for considering the perspectives of different sources. For example, dividing evidence into camps that are “for” or “against” an issue represents a simplifying heuristic for organizing and making sense of various stances on that issue.

A significantly stronger stance analysis involves research into, or leverages a preexisting knowledge of, the backgrounds and views of key individuals, groups, and thought leaders. Where appropriate, it may also include a more formal assessment that employs structured methods to identify critical relationships, to predict how the general worldview of a source is likely to influence his or her stance toward specific issues, or to detect the intentional manipulation of information. Social network analysis (Butts, 2008; Moon & Carley, 2007; Krebs, 2002), faction analysis (Feder, 1987; Smith, Branlat, Stephens, & Woods, 2008), and analysis of competing hypotheses (Heuer, 1999; Folker, 2000), for example, are techniques that may be selectively employed to those ends.

5. Sensitivity analysis. The term “sensitivity”, as it is used here, has a meaning most similar to its usage in the statistical analysis of quantitative variables, wherein it describes the extent to which changes in input parameters affect the output solution of a model (Saltelli, Chan, & Scott, 2000). However, rather than with the relationship between output and input variance, our concern is with the strength of an analytical assessment given the potential for low reliability and high uncertainty in supporting evidence and explanations. Phrased differently, sensitivity analysis describes the process of discovering the underlying assumptions, limitations, and scope of an analysis as a whole, rather than those of the supporting data in particular, as with the related attribute of information validation.

Many in the intelligence community emphasize the importance of examining analytical assumptions. Schum (1987), for example, identifies assumptions as foundational to analytic work. Similarly, Davis (2002) notes that “careful attention to selection and testing of key assumptions to deal with substantive uncertainty is now well established as the doctrinal standard for [analysis work]” (p. 5). As such, effective sensitivity analysis demands the
identification and explication of the key beliefs that tie an analysis together and that link evidence with hypotheses.

To that end, a strong sensitivity analysis goes beyond simple identification, meticulously considering the strength of explanations and assessments in the event that individual supporting evidence or hypotheses were to prove invalid. It also specifies the boundaries of applicability for the analysis. With weak sensitivity analysis, in contrast, explanations seem appropriate or valid at surface level, with little consideration of critical “what if” questions—e.g., “What if a key data source misidentified a person of interest?” Likewise, the overall scope of a weak analysis process may be unclear or undefined.

6. Information synthesis. Often emphasized by experts more than casual analysts is that rigorous analytical work is as much about putting concepts together as it is about breaking an issue apart (Cooper, 2005; Mangio & Wilkinson, 2008). That is to say, rigorous analysis demands not only “analytic” activity in the definitional sense, but “synthetic” activity as well. Thus, information synthesis is a reflection of the extent to which an analysis goes beyond simply collecting and listing data to provide insights not directly available in individual source data.

Weak information synthesis is reflected in analyses that succeed in compiling relevant and “on topic” information, but that do little in the way of identifying changes from historical trends or providing guidance for broader or more long-term concerns. Indicators of weak synthesis include extensive use of lists, copying material from other sources with little reinterpretation, and a lack of selectivity in what is emphasized by the analysis. A stronger synthesis is reflected by explicit efforts to develop an analysis within a broader framework of understanding. The depiction of events in relation to historical or theoretical context and the framing of key issues in terms of tradeoff dimensions and interactions also identify such analysis.

Stronger still is synthesis that has integrated information in terms of relationships rather than components, with a thorough consideration of diverse interpretations of relevant data. In addition, such synthesis is performed by reflexive analysts who are attentive to ways in which their particular analytical processes may hinder effective synthesis and who are attuned to the many potential “cognitive biases” that manifest in analytical work (Moore, 2007; Krizan, 1999; Heuer, 1999).

7. Specialist collaboration. Inevitably, analysts encounter topics on which they are not expert or that require multiple areas of expertise to fully
make sense of (Kent, 1966). Even in instances where an analyst has expertise in pertinent topics, success for the modern analyst still demands the incorporation of multiple perspectives on an issue (Clark, 2006; Medina, 2002). Accordingly, analytical rigor is enhanced when substantive expertise is brought to bear on an issue (Tenet, 1998).

The level of effort expended to incorporate relevant expertise defines effective specialist collaboration. In a process with little collaboration, minimal outside expertise is sought out directly. A moderately collaborative analysis process involves some interaction with experts, though at this level such expertise is often drawn from existing personal or professional networks, rather than from organizationally external sources. In a high-rigor process, independent experts in key content areas are identified and consulted. Thus, a strong specialist collaboration process is defined by efforts to go beyond a “core network” of contacts in seeking out domain-relevant expertise. In many cases, additional resources and “political capital” are expended to gain access to such specialized knowledge.

8. **Explanation critiquing.** Specialist collaboration and explanation critiquing are related in that both are forms of collaborative analytical activity that reflect the influence of diverse perspectives (Hong & Page, 2004; Guerlain et al., 1999). However, whereas specialist collaboration primarily relates to the integration of perspectives relative to information search and validation, explanation critiquing relates to the integration of perspectives relative to hypothesis exploration and information synthesis. More succinctly, explanation critiquing is concerned with the evaluation of the overall analytical reasoning process, rather than with the evaluation of content specifically.

Similar to specialist collaboration however, this attribute is largely defined by the extent to which analysts reach beyond immediate contacts in collecting and integrating alternative critiques. A low quality explanation critiquing process has limited instances of such integration, while a more moderate process leverages personal and professional contacts to examine analytical reasoning. In the latter case, it is often peer analysts, supervisors, or managers who serve as the primary source of these alternative critiques.

In a still stronger analysis process, independent as well as familiar reviewers have examined the chain of analytical reasoning and explicitly identified which inferences are stronger and which are weaker. In addition, the use of formal methods such as “red teams” or “devil’s advocacy” (Defense Science
Board Task Force, 2003; Pherson, 2005) as well as the expenditure of capital, political or otherwise, serve as hallmarks of strong explanation critiquing.

**Examples of Calibrating the Attributes of Rigor to a Particular Task**

The attribute-based framing of analytical rigor represents a model in the sense that it “constrains what can be measured by describing what is essential performance”; it represents an operational framework in that “the model parameters... become the basis for specifying the measurements” (Dekker & Hollnagel, 2004, p. 82). We now provide examples that illustrate how the attributes of rigor can be scored as low, moderate, and high for two very different analysis tasks.

The first example involves scoring the analytical rigor of twelve ad hoc, three-member teams that completed a logistics analysis and planning task set in a military context used in a prior study (Trent, Patterson, & Woods, 2007). The study task was for each team to develop a plan to move a predetermined number of troops and supplies to a desired location by the fastest, cheapest, and most secure routes possible, given a number of scenario-specific constraints. Half of the teams were collocated and worked together directly while the other half were distributed and limited to audio communication only.

For all cases, verbal transcripts of the study sessions were used to score the analysis and planning processes of the teams with respect to the attributes of rigor. Two independent coders rated each team as low, moderate, or high on each attribute as shown in Figure 2, with specialist collaboration not scored due to limitations imposed by the design of the study. Note that both coders were also provided with feedback as to how well the plan developed by each team met the stated scenario requirements.

Of the 84 items scored (12 teams x 7 scored attributes), 76 were judged by both raters to fall into the same low, moderate, or high category, implying strong agreement between raters ($\kappa_w = 0.86$). Disagreements in coding were most common with respect to the attribute of information search, but also occurred in scoring hypothesis exploration, information validation, and stance analysis. Overall however, there was general consistency in how the coders applied the framework to assess the analytical rigor of the processes employed by the teams.
Figure 2. Scoring the rigor of the analysis processes of teams in the first example, with differences between raters highlighted.

The second example involves another study in a laboratory environment in which participants were asked to respond to a quick reaction task in the form of a recorded verbal briefing that answered the question: “What were the causes and impacts of the failure of the maiden flight of the Ariane 501 rocket launch?” The task was conducted under conditions of data overload, time pressure, and low to no expertise in the area, and the scenario was based on the case of the Ariane 5 rocket, whose maiden launch on June 4, 1996, ended in a complete loss of the rocket booster and scientific payload when it exploded shortly after liftoff, an accident which was significant because it departed from historical launch failure patterns in its underlying causes.

Data originally collected by this study for other purposes were later reanalyzed with respect to the attributes of rigorous analysis for two junior intelligence analysts who had not yet received their security clearances (Miller & Patterson, 2008) and for eight senior intelligence analysts, all from the same intelligence agency (Patterson, Roth, & Woods, 2001). In this case, a single individual coded the processes of each analyst (see Figure 3), however note that neither specialist collaboration nor explanation critiquing were scored, as the study design did not allow participants to access individuals with specialized expertise or to have their problem framing and briefing critiqued by other analysts. The results of this evaluation are represented graphically in Figure 4.
### Table 3

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hypothesis</strong></td>
<td>No obvious initial bias.</td>
<td>Moderate</td>
<td>No obvious initial bias.</td>
<td>Moderate</td>
<td>Low</td>
<td>Generally sloppy, rushed to get an acceptable answer.</td>
<td>Moderate</td>
<td>Spent a lot of time working out the impacts, but focused on the diagnosis of the failure.</td>
<td>Moderate</td>
</tr>
<tr>
<td><strong>Exploration</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Information Validation</strong></td>
<td>Low</td>
<td>Wrong date for accident.</td>
<td>Moderate</td>
<td>Coordinated some information.</td>
<td>Low</td>
<td>Would relay on customer to clarify inconsistencies in data.</td>
<td>Moderate</td>
<td>Noticed discrepancies informally and kept track of when they came up again.</td>
<td>Low</td>
</tr>
<tr>
<td><strong>Stance</strong></td>
<td>Low</td>
<td>Malaysian MBA student treated the same as other sources.</td>
<td>Moderate</td>
<td>Sensitive when judging document contributions.</td>
<td>Moderate</td>
<td>Showed sensitivity to potential for bias in what was read, but did not obviously do anything when it was detected.</td>
<td>Moderate</td>
<td>Sensitive to stance when reading.</td>
<td>Moderate</td>
</tr>
<tr>
<td><strong>Analysis</strong></td>
<td>Low</td>
<td>No sensitivity to risk.</td>
<td>Low</td>
<td>No sensitivity to risk.</td>
<td>Moderate</td>
<td>Wanted to do timeline or ask customer to verify the right accident.</td>
<td>Moderate</td>
<td>Sensitive to risk.</td>
<td>Low</td>
</tr>
<tr>
<td><strong>Sensitivity</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Synthesis</strong></td>
<td>Low</td>
<td>Found document and forwarded as answer.</td>
<td>Low</td>
<td>Principal interests guided what he read next, with no obvious strategy for pulling together information.</td>
<td>Moderate</td>
<td>Provided insights that went beyond what was read, based on patterns and past experience.</td>
<td>Low</td>
<td>Briefing at level of what was written and not put in context.</td>
<td>Moderate</td>
</tr>
<tr>
<td><strong>Explanation Critiquing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Specialist Collaboration</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Critiquing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 4. Graphical representation of the rigor attribute scores shown in Figure 3.
Note that this depiction of the attributes affords an integrated, pattern-based assessment of rigor that reveals an overall picture of the relationships among the attributes, as well as insight into the level of rigor achieved on individual attributes with respect to the demands of a given context (Bennett & Flach, 1992). In this example, interestingly, stance analysis—and to a lesser extent information synthesis and sensitivity analysis—were most diagnostic of expertise on the study task. In considering both examples presented here, it is our observation that there is flexibility in the rigor attribute framework that allows for its application in a variety of analytical contexts and to a variety of assessment tasks.

**Discussion**

We have presented a framework for and examples of applying an approach to codify the rigor of an information analysis process from a macrocognitive perspective that reflects the diversity of strategies employed by professional analysts to cope with data overload, to make sense of an uncertain and dynamic world, and to reduce the risk of shallow analysis. In practice, however, scoring a given analytic process with respect to these eight attributes is not without challenges, and there are many opportunities for advancing this line of work.

Of principal concern is the fact that the rigor attribute model is grounded largely in the domain of intelligence analysis. Accordingly, the extent to which it proves relevant in other analytical domains remains an open issue, though preliminary research on activities such as information search by students working on educational projects, medical diagnosis using automated algorithms for x-ray interpretation, and accident investigation analyses (Patterson, Zelik, McNee, & Woods, 2008) suggest that the concepts may be fairly robust. Of course, the framework presented here does not purport to be the final word on rigor in information analysis either; rather, it represents a current iteration of an ongoing direction of exploration.

In contrast with the conventional, standards-based notions discussed earlier, our framework supports a view of analytical rigor that is better aligned with calls for guidance—as opposed to guidelines—in conducting analyses of complex and highly interconnected data (Patton, 2002). This is not to suggest that standard methods play an insignificant role in achieving rigor, but that such standards do not necessarily and completely define rigor for the practicing analyst. Rather it is the extent to which these methods and practices speak to the individual attributes that define the overall rigor of an
analysis. Of course, in the intelligence community in particular, this is hardly a surprising conclusion (Mangino & Wilkinson, 2008; Marrin, 2007; Johnston, 2005).

A related implication of the rigor attribute model is that, although it elucidates the aspects of analytical process that are perceived as most diagnostic of rigor, it does not directly reveal what is required to be “rigorous enough”. Again, this is not a surprising conclusion, though it is no less an important one.

In short, judging the sufficiency of rigor depends upon many contextual factors—analyst characteristics such as experience, specialization, and organizational affiliation; format and presentation characteristics such as writing style, presentation medium, and customer requirements; and production characteristics such as time pressure, resource availability, and priority, to name only a few (Krizan, 1999; Zelik, Woods, & Patterson, 2009). Moreover, just as there are situations where an analyst must rely on one or very few sources of information, so too are there situations in which exploring all available data is infeasible. Thus, contextual and situational demands largely dictate the best way in which to apply the rigor attribute model as a measure of process quality. Ultimately it is practitioners who negotiate and decide what is appropriate for a given set of circumstances (Gardiner, 1989; McLaughlin, 2008).

And so, in closing, we return to the risk of shallow analysis as a hallmark challenge of present-day information analysis. In the face of significant production pressures and rapidly proliferating data availability (Woods, Patterson, & Roth, 2002), it is increasingly easy for analysts and decision makers alike to fall into the trap of shallow analysis—believing an analysis to be sufficient in a context where it is not.

In responding to this challenge, the measurement of analytical rigor is advanced as a check against this risk. The model of analytical rigor presented here provides both direction for supporting this critical judgment task and a framework for a macrocognitive measure of performance that is based on how expert analysts assess rigor. To that end, the concept of analytical rigor in information analysis warrants continued exploration and diverse application as a macrocognitive measure of analytical sensemaking activity.
Acknowledgments

This research was supported by the Department of Defense (H98230-07-C-0406), the Air Force Research Laboratory (S110000012), and the Office of Naval Research (GRT0012190). The views expressed in this article are those of the authors and do not necessarily represent those of the Department of Defense, Air Force, or Navy. We thank Fernando Bernal for doing the analytic coding for the logistics analysis and planning task. We also thank David T. Moore, Richards J. Heuer, and Robert R. Hoffman for their insightful review comments.

References


T. Ormerod, & R. Lipshitz (Eds.), *Naturalistic decision making and macrocognition* (pp. 3–26). Aldershot, England: Ashgate.


