

The Supervisor's Dilemma: Judging When Analysis is Sufficiently Rigorous

Daniel Zelik, David D. Woods, Emily S. Patterson

The Ohio State University

1971 Neil Ave, Columbus, OH 43210

{zelik.2, woods.2, patterson.150}@osu.edu

Our perspective on sensemaking is grounded in the study of professional intelligence analysis. Though diverse in nature, the work of professional analysts accords with the description of sensemaking as a “motivated, continuous effort to understand connections (which can be among people, places, and events) in order to anticipate their trajectories and act effectively” (Klein, Moon, & Hoffman, 2006, p. 71) and observers have at times used the term “sensemaking” to describe all or part of analyst work (e.g., Pirolli, 2006; Mangio & Wilkinson, 2008). In fact, Cooper (2005, p. 42) notes the “primary purpose of analytic effort is ‘sensemaking’ and understanding, not producing reports; the objective of analysis is to provide information in a meaningful context, not individual factoids.”

Our research explores sensemaking in this domain by studying how professional intelligence analysts assess the quality of their analysis products—products which represent an effort to “understand complex, dynamic, evolving situations that are ‘rich with various meanings’” (Hutchins, Pirolli, Card, 2006). Specifically, we explore sensemaking relative to the question of how analysts judge analytical rigor, through a critiquing study that frames the assessment of rigor in terms of a critical judgment task we describe as the “Supervisor’s Dilemma” (Zelik, Patterson, & Woods, 2007a).

The Supervisor’s Dilemma describes a generic situation wherein a supervisor must decide if the output product of an analysis is acceptably rigorous or if more analytical resources must be invested in that analysis process before sending it forward. This judgment provides a mechanism for exploring the understanding of rigor in analysis because, while in principle it is an abstraction of a common occurrence, it represents a critical decision that, in practice, is often made tacitly (Betts, 1978). Its value for studying sensemaking, however, stems primarily from the character

of the dilemma itself. In this dilemma, achieving “maximum” rigor is not an option; rather, it is about monitoring and adjusting the process of analysis relative to the demands of a dynamic world and about being rigorous enough. Thus, the Supervisor’s Dilemma effectively captures the fundamental and ongoing tradeoffs inherent in analysis work—balancing limited resources (time, expertise, sources, etc.) against the need to resolve uncertainty, balancing production pressures against the need to broaden the analytical process.

Framing the Supervisor’s Dilemma as a generic judgment task also circumvents some of the controversies within the intelligence community surrounding the role of decision makers relative to the analysis process. Structuring this critical judgment as a dilemma of the supervisor frames the decision as one that is simultaneously explicit, collaborative, and yet largely informal. On one hand, the dilemma is not so personal that the decision is made implicitly by an analyst and yet not so formal that the analysis is used by a decision maker as, perhaps, the sole basis for a policy decision.

Thus, assuming the role of supervisor reflects a perspective somewhere between that of analyst and that of decision maker, providing an agreeable point on the continuum for eliciting feedback from study participants. As such, it embodies many related questions that an analyst, supervisor, or decision maker might ask about an analysis product—e.g., is the analysis finished? Is the analysis ready to send forward? Is the analysis good enough to act on? If not, what else needs to be known? Given what is unknown, what is the best way to invest limited analytical resources to reduce this uncertainty and improve the analysis product?

The Supervisor’s Dilemma, however, represents more than just a series of hypothetical questions. The dilemma highlights the criticality of the interactions that occur at the interfaces between individuals within an intelligence organization. It embodies a point in the intelligence cycle where a decision is made—whether by supervisors, decision makers, or the analysts themselves—as to whether or not an analysis product is ready to pass on to the next stage in the cycle. And, it represents a point where an individual or group’s understanding of a circumstance “is turned into words and salient categories... embodied in written and spoken texts,” and shared with others (Weick,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Copyright is held by the author/owner(s).

CHI 2009, April 4–9, 2009, Boston, Massachusetts, USA.

ACM 978-1-60558-247-4/09/04...\$5.00

Sutcliffe, & Obstfeld, 2005). Simply, whether made implicitly or explicitly, the Supervisor's Dilemma addresses a fundamental question in all forms of information analysis—when is an analysis sufficient?

To that end, the Supervisor's Dilemma makes explicit the continuous demand that professional analysts face to balance confidence with cautiousness, as "wise" analysts are ever-cognizant "that they don't fully understand what is happening... because they have never seen precisely this event before" (Weick, 1993). Thus, the dilemma also highlights the risk, inherent in all information analysis processes, that an analysis is of inadequate depth relative to the demands of the situation—a risk we characterize as the

"risk of shallow analysis" (Zelik, Patterson, & Woods, 2007a). To fall into the analytical trap of overconfidence in inadequate analysis—or, conversely, of under-confidence in adept analysis—is to risk a failure of analytical sensemaking. A vivid example of this vulnerability was revealed by the accident investigation that followed the loss of the shuttle Columbia, where reports revealed critical incidents in which decisions were made based on analyses that appeared to be thorough, but that were in fact of very low rigor (Columbia Accident Investigation Board, 2003; Woods, 2005).

In addressing the risk of shallowness within the context of intelligence analysis, critically evaluating the underlying

Attribute	Description
Hypothesis Exploration	Hypothesis Exploration describes the extent to which multiple hypotheses were considered in explaining data. In a low-rigor process there is minimal weighing of alternatives. A high-rigor process, in contrast, involves broadening of the hypothesis set beyond an initial framing and incorporating multiple perspectives to identify the best, most probable explanations.
Information Search	Information Search relates to the depth and breadth of the search process used in collecting data. A low-rigor analysis process does not go beyond routine and readily available data sources, whereas a high-rigor process attempts to exhaustively explore all data potentially available in the relevant sample space.
Information Validation	Information Validation details the level at which information sources are corroborated and cross-validated. In a low-rigor process little effort is made to use converging evidence to verify source accuracy, while a high-rigor process includes a systematic approach for verifying information and, when possible, ensures the use of sources closest to the areas of interest.
Stance Analysis	Stance Analysis is the evaluation of data with the goal of identifying the stance or perspective of the source and placing it into a broader context of understanding. At the low-rigor level an analyst may notice a clear bias in a source, while a high-rigor process involves research into source backgrounds with the intent of gaining a more subtle understanding of how their perspective might influence their stance toward analysis-relevant issues.
Sensitivity Analysis	Sensitivity Analysis considers the extent to which the analyst considers and understands the assumptions and limitations of their analysis. In a low-rigor process explanations seem appropriate and valid on a surface level. In a high-rigor process the analyst employs a strategy to consider the strength of explanations if individual supporting sources were to prove invalid.
Specialist Collaboration	Specialist Collaboration describes the degree to which an analyst incorporates the perspectives of domain experts into their assessments. In a low-rigor process little effort is made to seek out such expertise, while in a high-rigor process the analyst has talked to, or may be, a leading expert in the key content areas of the analysis.
Information Synthesis	Information Synthesis refers to how far beyond simply collecting and listing data an analyst went in their process. In the low-rigor process an analyst simply compiles the relevant information in a unified form, whereas a high-rigor process has extracted and integrated information with a thorough consideration of diverse interpretations of relevant data.
Explanation Critique	Explanation Critique is a different form of collaboration that captures how many different perspectives were incorporated in examining the primary hypotheses. In a low-rigor process, there is little use of other analysts to give input on explanation quality. In a high-rigor process peers and experts have examined the chain of reasoning and explicitly identified which inferences are stronger and which are weaker.

Table 1. Attributes of rigorous analysis (Adapted from Zelik, Patterson, & Woods, 2007a)

rigor of an analysis product represents an approach that professional analysts employ to cope with this fundamental uncertainty and to calibrate their level of understanding of events in the world. Thus, assessing rigor is viewed as an expert behavior that contributes to effective sensemaking in intelligence analysis.

Our research identified two general findings on rigor in intelligence analysis. First, it suggests a model of analytical rigor that frames the concept as an emergent multi-attribute measure of sufficiency ("Has the analyst done enough?") rather than as it is often conventionally framed, as a measure of process adherence ("How rigidly has the analyst followed a particular method?"). Second, it reveals that insight into the analysis process influences both perceptions of and judgments about the quality of an analysis product (Zelik, Patterson, & Woods, 2007c).

Regarding the first finding, the rigor attribute model represents a model of sensemaking behavior that emerged from the study of practicing analysts, which frames the concept of rigor as the composite of multiple process traits. This multi-attribute model characterizes these indicators as independent aspects of the analysis process which, when aggregated, reveal a composite assessment of analytical rigor. That is to say, in as much as there is rigor in an analysis process, these are the cues to which the analysts were attuned in making their assessments. Specifically, the study identified eight critical attributes of analysis processes that contribute to assessments of rigor (See Table 1).

Regarding the second finding, our research suggests that supporting the judgment of analytical rigor is about revealing insight into the analysis process. Tufte (2003), for example, is critical of the influence of tools like PowerPoint on the understanding of analytical process, noting that such software contributed to the acceptance of unacceptably low-rigor analyses in the Columbia case. Our research offers

direction for coping with this challenge of representing and sharing analysis. Thus, this research has applied value in supporting the distributed sensemaking that is the hallmark of the intelligence community—in particular, in supporting those interactions which occur between analysts and policy makers. Current efforts focus on applying the model to aid sensemaking via the visualization of analysis processes in a way that supports the assessment of rigor (From Zelik, Patterson, & Woods, 2007b). Figure 1 illustrates one of many possible attribute-based representations of an analysis process.

The measure of success in resolving the Supervisor's Dilemma is the level of discrimination in the sufficiency judgement—the ability to recognize both when and where to invest additional efforts into improving an analysis process and a judgment made not once but iteratively against a backdrop of change. Innovating new and more effective forms of representing the analysis process, grounded in an understanding of how professional analysts judge analytic rigor, offers direction for making the sufficient rigor judgement—for avoiding the trap of shallow analysis, for developing tools and methods that convey a meaningful understanding of analytical process, and for supporting sensemaking across analysis contexts. Perhaps most importantly though, this research into rigor via the Supervisor's Dilemma reveals promising directions for the continued exploration of the challenges—those long-standing as well as those driven by recent technological change—that confound sensemaking in intelligence analysis.

ACKNOWLEDGMENTS

This research was supported by the Department of Defense (BAA-001-04). The views expressed are those of the authors and do not necessarily represent the view of the Department of Defense.

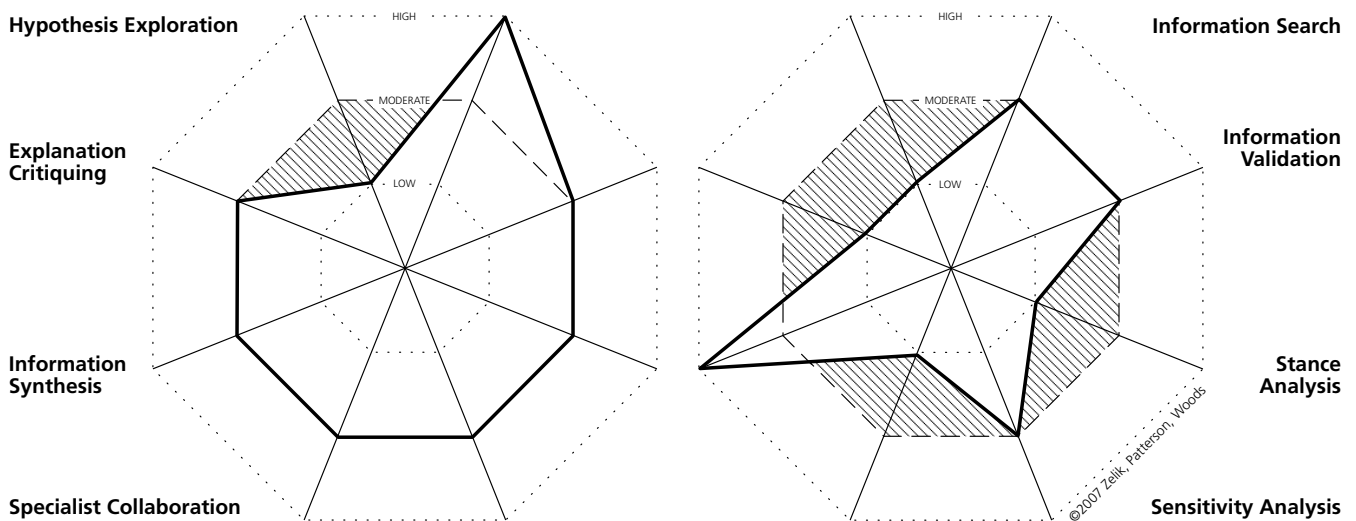


Figure 1. One representation of an attribute-based framing of rigor. (Adapted from Zelik, Patterson, & Woods, 2007a)

REFERENCES

1. Betts, Richard K. (1978). Analysis, war, and decision: Why intelligence failures are inevitable. *World Politics*, 31, 61–89.
2. Columbia Accident Investigation Board. (2003). *Columbia Accident Investigation Board report*. Washington, DC: U.S. Government Printing Office.
3. Hutchins, S. G., Pirolli, P. L., & Card, S. K. (2006). What makes intelligence analysis difficult? A cognitive task analysis of intelligence analysts. In P. L. Pirolli (Ed.), *Assisting people to become independent learners in the analysis of intelligence* (N00014-02-C-0203, pp. 6–52). Palo Alto, CA: Palo Alto Research Center, Inc.
4. Klein, G., Moon, B., & Hoffman, R. R. (2006). Making sense of sensemaking 1: Alternative perspectives. *Intelligent Systems*, 21, 70–73.
5. Mangio, C. A., & Wilinson, B. J. (2008, March). *Intelligence Analysis: Once Again*. Paper presented at the International Studies Association 2008 Annual Convention, San Francisco, CA.
6. Pirolli, Peter. (2006). *Assisting People to Become Independent Learners in the Analysis of Intelligence* (N00014-02-C-0203). Palo Alto, CA: Palo Alto Research Center, Inc.
7. Tufte, Edward R. (2003). *The cognitive style of PowerPoint*. Cheshire, CT: Graphics Press LLC.
8. Weick, Karl E. (1993). The collapse of sensemaking in organizations: *The Mann Gulch disaster*, 38, 628–652.
9. Weick, K. E., Sutcliffe, K. M., & Obstfeld D. (2005). Organizing and the process of sensemaking. *Organization Science*, 16, 409–421.
10. Woods, D. D. (2005). Creating foresight: Lessons for resilience from Columbia. In W. H. Starbuck and M. Farjoun (Eds.), *Organizing at the limit: NASA and the Columbia disaster*. Malden, MA: Blackwell.
11. Zelik, D., Patterson, E. S., & Woods, D. D. (2007a, June). *Understanding rigor in information analysis*. Paper presented at the 8th International Conference on Naturalistic Decision Making, Pacific Grove, CA.
12. Zelik, D., Patterson, E. S., & Woods, D. D. (2007b, August). *Supporting the assessment of rigor: Representing analysis to create process insight*. Paper presented at the 2nd Annual Workshop on Meta-Information Portrayal, Washington, DC.
13. Zelik, D., Patterson, E. S., & Woods, D. D. (2007c, October). *Judging sufficiency: How professional intelligence analysts assess analytical rigor*. Paper presented at the Human Factors and Ergonomics Society 51st Annual Meeting, Baltimore, MD.