

Judging Sufficiency: How Professional Intelligence Analysts Assess Analytical Rigor

Daniel J. Zelik
The Ohio State University
Columbus, Ohio

Emily S. Patterson
The Ohio State University
Columbus, Ohio

David D. Woods
The Ohio State University
Columbus, Ohio

This study examines how professional intelligence analysts judge the rigor behind an analysis. The study investigates the challenges that inhibit the understanding of rigor in intelligence analysis and explores cues used by analysts to identify analytic rigor—or lack of rigor. Nine professional intelligence analysts participated in a modified elicitation by critiquing method study, embedded in a scenario walkthrough. Findings from the study indicate that, while professional intelligence analysts can make perceptive assessments about the quality of an analysis process based on product quality, these perceptions are apt to change with insight into the analytic process.

INTRODUCTION

Recent events prompted many within the U.S. intelligence community to question the present state of the professional intelligence analysis tradecraft (National Commission on Terrorist Attacks, 2004; Duelfer, 2004; Johnson, 2005). In the ensuing debates, it became apparent that changes in technology and in the international political landscape—among other influences—have transformed the role of the professional analyst (Medina, 2002; Ward, 2002). In response, efforts are underway to help analysts (e.g., the ARDA NIMD program) adapt to this changing environment (Pirolli, 2006; Greitzer, 2005; Elm, et al., 2005). Such research aims to improve the analysis process by better understanding the critical interactions that influence performance.

This study asks how analysts decide when there is “sufficient rigor” in an analytic process. Rigor, as an assessment of process quality, is used in information analysis to communicate about the process, rather than the product, of analysis (Zelik, Patterson, & Woods, 2007). Our research on rigor in information analysis began in engineering safety analysis with a case where accident reports revealed that managers were unaware they were making decisions based on analyses that appeared thorough on the surface, but were in fact of very low rigor (Columbia Accident Investigation Board [CAIB], 2003; Crippen, et al., 2005).

This research serves to expand the understanding of rigor by studying it in the context of intelligence analysis. Specifically, it explores cues analysts use to identify analytic rigor—or lack of rigor—and investigates the challenges that inhibit the understanding of rigor in intelligence analysis. Our study used an innovative approach, based upon a critiquing knowledge elicitation methodology, to tap into the domain expertise of the professional intelligence analyst.

METHOD

Participants

Nine professional intelligence analysts participated in the study, consisting of eight primary participants and one pilot study participant. The pilot participant was distinct from the primary participants in that he represented a different intelligence organization than the other eight. All professional analysts who participated in the study represented the strategic, in contrast with the tactical, intelligence analysis community. The study included analysts with a broad range of

professional experience (*Mdn* = 20 years, *M* = 21 years, *SD* = 10 years), from junior analyst with less than three years of experience to senior analyst with over forty years of experience. The participants (P0–P8) also varied with respect to their intelligence backgrounds, current principle duties (e.g., professional analyst, linguist, analyst supervisor, etc.), and gender (6 male, 3 female). The analysts selected for the study were volunteers who responded to an e-mail request for participants via a contact at the primary intelligence organization. No compensation was provided for participation.

Scenario

The study scenario focused on the importation of Liquefied Natural Gas (LNG) within the U.S. energy sector. LNG importation is a relevant issue in many regions of the United States, given the current state of U.S. energy needs relative to the concerns surrounding homeland safety and security (Clarke, 2005). This topic, selected in consultations with professional analysts and an industry expert, was chosen based on its primary attributes of authenticity, unclassified domain information, scope, and topical interest.

In the scenario, two hypothetical junior intelligence analysts—Analyst 1 and Analyst 2—were tasked by a decision maker to answer a question related to LNG. The task question driving the scenario was “What are the primary obstacles to using Liquefied Natural Gas to address energy needs in the U.S.?” The analytic responses of Analyst 1 and Analyst 2 were the basis for critique in the study—Analysis Process 1 [A(1)] and Analysis Process 2 [A(2)], respectively.

The task question was framed as a quick reaction task, meaning that the junior analysts each spent a similar amount of time producing a response to the question. The backgrounds of the analysts were also similar, so as not to be a significant factor in the scenario. Throughout the study, participants assumed the role of the hypothetical supervisor of these junior analysts, responsible for deciding if and when the analysis briefings were of acceptable quality to forward to the decision maker who initiated the task question request.

In designing the scenario, the briefing report of Analyst 1 was based on an analytical response developed by the researchers, while the report of Analyst 2 was based heavily on an open source strategic forecast on the future of LNG in the U.S. Each briefing was two pages in length. Because the study was scenario-based, these analytic responses used were not taken directly from any real case. Rather, they were representative analyses designed to amplify

the cues that it was initially hypothesized that analysts use to infer rigor. A(1) was designed to reflect a high-rigor process, while A(2) was designated as a contrasting low-rigor process.

Procedure

The design of this study was based on the Elicitation by Critique (EBC) methodology (Miller, Patterson, & Woods, 2006). In this approach, participants share their expertise by critiquing the processes of other domain practitioners, rather than by directly verbalizing individual experiences. The approach used in this study was modified from the original EBC methodology in three ways. First, participants critiqued not only the processes of other analysts, but also the briefing products produced by those processes. Second, participants critiqued the analyses of two different analysts concurrently, rather than sequentially. Third, participants critiqued analyses that were generated to fit within the context of a scenario. Because of these modifications from the original method, the design of this study was a modified EBC methodology, embedded in a scenario walkthrough.

Each participant engaged in a two-part critique during a single elicitation session. In the first part of the study, critiquing focused on the analysis products, while in the second it focused on the analysis processes. The sessions also included an introductory overview and a concluding debrief.

During the initial critique cycle—focused on the analysis product—participants were presented with the two analysis briefings generated via A(1) and A(2). The order in which the briefings were presented was randomized across sessions. Participants were instructed to review both reports and "assess and compare the quality of the two written reports" and "comment on the 'rigor' of the processes that produced each of the reports." Participants were then given up to fifteen minutes to individually review the two documents. During this time, researchers left each participant alone while they assessed the reports.

After participants completed an initial review of the documents, they signaled for the researchers to return. At this time, the researchers prompted participants to verbally critique the two briefings (1) by comparing and contrasting them, (2) by commenting on what they could infer about the rigor of the process that produced each report, and (3) by describing the approach they would use to better understand the rigor behind each report. Participant feedback was collected as verbal report data. This critique cycle concluded when participants were asked, in their role as supervisor within the scenario, to decide if either, both, or neither of the analysis briefings were ready to be forwarded on to the decision maker—and thus putting themselves into the *Supervisor's Dilemma*.

The *Supervisor's Dilemma* describes a generic situation wherein a supervisor must decide if the output product of an analyst is acceptably rigorous or if more resources must be invested to improve the analysis before sending it forward. The dilemma facing the study participants, as proxy supervisors, was to decide if the scenario-based analysis product was ready to forward on to a decision maker by assessing the rigor of the process that produced it.

The second critique cycle—focused on assessing the analysis process—began when participants were presented with a list of seven process documents, listed in Table 1, that could be used to assess the analysis processes. Participants

were asked to comment on the perceived value of having access to each of the process documents in making an assessment of process rigor, ranking the seven options. In addition to verbal report data, ordinal preference rankings were collected.

Table 1. Process Documents Available to Participants

Document Type	Analysis 1	Analysis 2
Query Summary	4 query targets	3 query targets
Documents Read (in order of reading)	26 documents	7 documents
Where Documents Stored in Folders	7 folders	4 folders
Key Documents (relied on heavily)	4 key documents	3 key documents
Hypotheses – Considered Obstacles to LNG	5 hypotheses	3 hypotheses
Collaborations with Others During Process	6 collaborations	1 collaboration
Research Note-Sheet	Yes	No

Next, participants were given access to two sets of process documents, reflecting A(1) and A(2). These documents were used by participants to construct an understanding of the analytic processes that produced the reports. Participants were again given up to fifteen minutes to review the process documents. After examining these documents and signaling for the researchers to return, participants were asked to critique the analysis processes (1) by comparing and contrasting the processes, (2) by commenting on the rigor they saw in each process, and (3) by comparing their revised assessments of rigor to their earlier perceived assessments of rigor. Feedback was collected as verbal report data. The second critique cycle concluded with a revisiting of the *Supervisor's Dilemma*, with participants now having a more informed understanding of the rigor of the analysis processes.

Each of the elicitation sessions was conducted with participants individually, with either one or two researchers present. Each session was scheduled for one and a half hours, with the exception of P8, whose session was scheduled for one hour due to time constraints. The study sessions were run over four days at times convenient for the participants.

Data Analysis

The data collected in the research study were analyzed by organizing the participant feedback relative to four elicitation prompts designed into the study. The study included two decision prompts—rigor assessment and *Supervisor's Dilemma*—that spanned multiple phases of the study and that were designed to capture how process insight influenced perceptions of rigor. The study also included two discussion prompts, which corresponded to the two critiquing cycles described in the procedure and served to elicit responses from participants about the cues they used to infer rigor relative to comparisons of the briefing reports and process documents. Each of these two discussion prompts also included a number of embedded probes, designed to push the scenario beyond a standard case (Woods & Hollnagel, 2006).

Rigor Assessment Prompt. This prompt asked analysts to decide which analysis process they believe to be more rigorous, both before and after seeing the process documents. It was targeted at understanding what resources, or process documents, analysts used to make assessments about rigor. This prompt required participants to quickly shift from thinking abstractly (i.e., thinking about what they might want to use to assess process quality) to thinking very concretely about the process documents relative to the scenario.

Supervisor's Dilemma Prompt. This prompt spanned across the phases of the study, as participants were asked to make an assessment of analysis report readiness before and after viewing the process documents. In doing so, participants were confronted with a situation wherein the new information has either supported or contradicted their initial assessments about the quality of the analysis products. Their attempts to resolve this potential inconsistency provides insight into their understanding of how product quality relates to process rigor.

Comparison of Analysis Reports Prompt. At the outset of the study, participants understood that the two versions of the analysis briefing were created by two similar analysts working the same questions under similar constraints. Yet, the two reports differed significantly in their responses to the tasked question, forcing participants to resolve this discrepant information. Resolving these differences operationally defined the prompt within the context of the study. It was driven by the interaction of three embedded probes, listed in Table 2.

Table 2. Comparison of Analysis Report, Embedded Probes

Sub-Probe	Description
Tone of Reports	The two briefings conveyed distinctly different tones, revealing a difference in perspective between author analysts. Analyst 2 seemed to favor LNG as a viable solution to increasing U.S. energy need, while Analyst 1 did not seem to support this position. In addition, the briefing reports differed in tone relative to writing style as well as in the subtle differences in briefing format and organization.
Obstacle Set	The difference in author positions also manifested in the obstacles that were cited in each report as those primary in prohibiting the more widespread acceptance of LNG. The reports had only partially overlapping obstacle sets, which probed participants to push further into each report, rather than accepting each as being complete at face value.
Inclusion of Figure	The A(2) report included a line graph related to LNG, while A(1) did not. This graphic prompted participants to comment on the distinction between the quality of content of each report and the presentation qualities of each report. Another dimension added by the inclusion of the graphic stemmed from the fact that the figure was not directly relevant to the primary analysis question, as it was disconnected from the core arguments presented in the A(2) briefing.

Comparison of Analysis Processes Prompt. In the study, the two processes were designed to be distinctly different. One analysis process was performed via what was characterized as a high-rigor process, while the other was performed via a low-rigor process. This difference in process forced participants to reinterpret prior assessments of report quality relative to the new understanding that was revealed by reviewing process information. There were seven unique process documents for each report that were included in the study, each of which also represented an embedded probe as described in Table 3.

Table 3. Comparison of Analysis Process, Embedded Probes

Sub-Probe	Description
Number of Documents Read	A(1) and A(2) had a nearly four-fold difference in number of documents read. The number of documents read reflected both on the overall rigor of the analytic process, but also on the experience and knowledgeability of the author analysts.
Number of Collaborations	A(1) showed a substantially greater incorporation of alternative perspectives through collaboration. It included both more and better quality collaborations than A(2).
Organization of Documents by Folder	The two processes differed on the extent to which documents were organized during the analysis process and also in the approach used to organize the document sets. The A(1) documents were better organized.
Number of Hypotheses Considered	A(1) reflected a process in which more hypotheses were weighed and considered than in A(2). Additionally, the nature of the hypotheses that were investigated varied between the two processes.
Nature of Key Documents	The type of document identified as key differed between processes. In A(1) the key documents were classified as formal research reports whereas A(2) relied on a varied mix of less formal supporting documents.
Inclusion of Research Note-Sheet	For A(1) study participants were able to view the collection of notes used in preparing the final briefing report. In contrast, Analyst 2 did not use a process that included a note-sheet. This probe offered the participant a number of options in explaining this discrepancy.
Type of Queries	While the total number of queries did not differ substantially between the processes, query type did. A(1) used more targeting queries and broader queries, while A(2) tended to use queries focused on a main hypothesis.

Note. Sub-probes in Table 3 correspond to a process document in Table 1.

RESULTS

The data indicate that providing insight into an analysis process influences assessments of both process rigor and product quality (Table 4). With respect to the rigor assessment prompt, five of nine participants, after reviewing the process documents, altered their assessments of whether A(1) or A(2) was more rigorous. The data also indicated that neither analysis process was perceived as substantially more rigorous than the other, as two of nine participants identified A(2) as being more rigorous, even after viewing the process documents.

Table 4. Run Order, Rigor Assessment, and *Supervisor's Dilemma*, Before and After Reviewing Process Documents

ID	Order	Rigor Assessment		Supervisor's Dilemma	
		Before	After	Before	After
P0	A(1)	A(2)	A(1)	A(2)	A(2)
P1	A(1)	A(1)	A(2)	A(1)	–
P2	A(1)	A(2)	A(1)	–	–
P3	A(2)	A(1)	A(1)	–	–
P4	A(1)	A(1)	A(1)	–	–
P5	A(2)	A(1)	A(1)	–	–
P6	A(2)	A(1)	A(1)	–	–
P7	A(2)	A(1)	A(2)	A(1)	–
P8	A(1)	A(2)	A(1)	–	–

This result also extended to the *Supervisor's Dilemma* probe, as relatively few participants felt either reports was ready to send to a decision maker. Before seeing the process

DISCUSSION

While participants made insightful judgments about the quality of an analytic process based on product quality, the findings of this study indicate that these perceptions were apt to change with the addition of process insight. This highlights an important distinction between *perceived* rigor, based on cues inferred from an analytic product, and *effective* rigor, based on insight into the analytic process—indicating that they often may not be aligned. Moreover, it was observed that even partial, incomplete, and second-hand process information initiated a change in understanding, as assessments of rigor in the study changed based on seeing process documents, rather than on the direct observation of an analysis process.

This finding both supports and extends the observation of Tufte (2003) that the form of an analysis product—in this case a briefing report, rather than a PowerPoint—can distort the understanding of the process that produced it. But, it is important to recognize that conveying an understanding of an analysis process is not simply revealing information about that process. While many of the same qualities of an analytic product were used to make inferences about analytic process, the study results show substantial differences in how conclusions were drawn from these cues.

This finding echoes an observation shared by a number of participants in the study—that there is no one right way to do analysis. Rather, one must accept that others have different practices that produce acceptable results—and that their approaches can be equally valid, even if not completely understood. This perspective on rigor in the analysis process extends beyond the intelligence analysis community into other forms of information analysis. The importance of effectively communicating analytical rigor was brought to the forefront of NASA's focus following the Columbia accident (Crippen, et al., 2005; CAIB, 2003). Private industry is also increasingly recognizing the value of understanding the appropriate level of rigor in information analysis processes relative to decision criticality and resource priority (Krizan, 1999). Thus, studying rigor offers insight into the difficulties that frustrate analysis processes across domains.

Although this study of rigor is restricted in its generalizability, given pragmatic limitations on sample size and study design, the findings do serve to further the understanding of rigor. Given current analytical production pressures (Johnson, 2005; Heuer, 1999) and the technology-driven proliferation of data availability (Patterson, Roth, & Woods, 2001) shaping the intelligence community, it is increasingly difficult to accurately judge the sufficiency of an analysis. A better understanding of rigor represents an analytic broadening check to be leveraged against this constant risk (Zelik, Patterson, & Woods, 2007). Perhaps more importantly though, this research into analytical rigor reveals promising directions for the continued exploration of the challenges faced by the professional analyst.

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.

REFERENCES

- Clarke, Richard A. (2005, May). *LNG facilities in urban areas* (GHC-RI-0505A). Arlington, VA: Good Harbor Consulting, LLC.
- Columbia Accident Investigation Board. (2003). *Columbia Accident Investigation Board report*. Washington, DC: U.S. Government Printing Office.
- Crippen, D. L., Daniel, C. C., Donahue, A. K., Helms, S. J., Livingstone, S. M., O'Leary, R., & Wegner, W. (2005). Annex A.2: Observations by Dr. Dan L. Crippen, Dr. Charles C. Daniel, Dr. Amy K. Donahue, Col. Susan J. Helms, Ms. Susan Morrisey Livingstone, Dr. Rosemary O'Leary, and Mr. William Wegner. In Return to Flight Task Group (Eds.), *Return to Flight Task Group final report* (pp. 188–207). Washington, DC: U.S. Government Printing Office.
- Duelfer, Charles. (2004). *Comprehensive report of the Special Advisor to the DCI on Iraq's WMD*. Washington, DC: U.S. Government Printing Office.
- Elm, W., Potter, S., Tittle, J., Woods, D. D., Grossman, J., & Patterson, E. S. (2005). Finding decision support requirements for effective intelligence analysis tools. In *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.
- Greitzer, Frank L. (2005, March). *Methodology, metrics, and measures for testing and evaluation of intelligence analysis tools* (PNWD-3550). Richland, WA: Battelle Memorial Institute, Pacific Northwest Division.
- Heuer, Richards J. (1999). *Psychology of intelligence analysis*. Washington, DC: U.S. Government Printing Office.
- Johnson, Rob. (2005). *Analytic culture in the US Intelligence Community: An ethnographic study*. Washington, DC: U.S. Government Printing Office.
- Krizan, Lisa (1999, June). *Intelligence essentials for everyone* (Joint Military Intelligence College Occasional Paper Number Six). Washington, DC: U.S. Government Printing Office.
- Medina, Carmen A. (2002). The coming revolution in intelligence analysis: What to do when traditional models fail. *Studies in Intelligence: Journal of the American Intelligence Professional*, 46, Article 03. Retrieved February 10, 2007, from <https://www.cia.gov/csi/studies/vol46no3/article03.html>
- Miller, J. E., Patterson, E. S., & Woods, D. D. (2006). Elicitation by critiquing as a cognitive task analysis methodology. *Cognition, Technology & Work*, 8, 90–102.
- National Commission on Terrorist Attacks. (2004). *The 9/11 commission report: final report of the National Commission on Terrorist Attacks upon the United States*. Washington, DC: U.S. Government Printing Office.
- Patterson, E. S., Roth, E. M., & Woods, D. D. (2001). Predicting vulnerabilities in computer-supported inferential analysis under data overload. *Cognition, Technology & Work*, 3, 224-237.
- 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.
- Tufte, Edward R. (2003). *The cognitive style of PowerPoint*. Cheshire, CT: Graphics Press LLC.
- Ward, Steven. (2002). Counterpoint to "The coming revolution in intelligence analysis": Evolution beats revolution in analysis. *Studies in Intelligence: Journal of the American Intelligence Professional*, 46, Article 04. Retrieved February 10, 2007, from <https://www.cia.gov/csi/studies/vol46no3/article04.html>
- Woods, D. D., & Hollnagel, E. (2006). *Joint cognitive systems: Patterns in cognitive systems engineering*. New York: Taylor & Francis.
- Zelik, D., Patterson, E. S., & Woods, D. D. (2007, June). *Understanding Rigor in Information Analysis*. Paper presented at the 8th International Conference on Naturalistic Decision Making, Pacific Grove, CA.