PNWD-SA-6893

# Instrumenting the Intelligence Analysis Process

Ernest Hampson
Paula Cowley

Presented at:

First International Conference on Intelligence Analysis
Methods and Tools
2-6 May, 2005
MITRE, McLean, VA, USA


**Battelle Memorial Institute**

# Instrumenting the Intelligence Analysis Process

Ernest Hampson
*Battelle, Colonial Place Operations*
*2111 Wilson Blvd., Suite 1000*
*Arlington, VA 22201*
*hampsone@battelle.org*

Paula Cowley
*Battelle, Pacific Northwest Division*
*3350 Q Avenue, K7-22*
*Richland, WA 99352*
*paula.cowley@pnl.gov*

**Keywords:** Novel Intelligence from Massive Data, instrumentation, analytic methods and tools, analysis process

## Abstract

The Advanced Research and Development Activity initiated the Novel Intelligence from Massive Data (NIMD) program to develop advanced analytic technologies and methodologies. In order to support this objective, researchers and developers need to understand what analysts do and how they do it. In the past, this knowledge generally was acquired through subjective feedback from analysts. NIMD established the innovative Glass Box Analysis (GBA) Project to instrument a live intelligence mission and unobtrusively capture and objectively study the analysis process. Instrumenting the analysis process requires tailor-made software hooks that grab data from a myriad of disparate application operations and feed into a complex relational database and hierarchical file store to collect, store, retrieve, and distribute analytic data in a manner that maximizes researchers' understanding. A key to success is determining the correct data to collect and aggregating the low-level data into meaningful analytic events. This paper will examine how the GBA team solved some of these challenges, continues to address others, and supports a growing user community in establishing their own GBA environments and/or studying the data generated by GBA analysts working in the Glass Box.

## 1. Introduction

The Glass Box is the hub project of NIMD and provides a software environment for the collection of data on the intelligence analysis (IA) process and facilities for the examination of the collected data. It is also a software integration platform for other NIMD research tools. NIMD researchers can interpret implicitly from the captured data and explicitly from the analysts' annotations the meaning behind the analysts' actions and their influence on the IA process.

A challenge since the beginning of the Glass Box project has been to determine which data elements should be collected for the researchers, normally before the researchers themselves knew which data they would require. Gorton and Haack (Gorton and Haack, 2004) documented this challenge as it relates to the software engineering component of the Glass Box development. However, separate from the difficulties of implementing an application without concrete requirements, there is the cognitive problem of determining which data is important and reflective of the analytic process. Capturing data on the analytic process in the Glass Box data will hopefully support insights into the NIMD primary research areas: modeling analysts and analytic processes, capturing and reusing prior and tacit knowledge, generating and managing hypotheses, and human information interaction. Although we have found that much of this data can be captured on-line from the analysts' workstation activities, much of the analytic process occurs off line - in the mind of the analyst, in collaborative conversations with other analysts, and in odd moments of reflection while doing other tasks (Cowley, *et al.* 2005).

## 2. Glass Box Analysis

According to the NIMD home page (http://www.ic-arda.org/Novel_Intelligence/index.html), "*The central focus of **NIMD** is on analysts and how they work, and more specifically on supporting analysts as the orchestrators and directors of the analytic process*". From the very beginning, the Glass Box project included the participation of professional intelligence analysts who work in the Glass Box and are highly motivated to inform and foster the development of more effective intelligence analysis tools. The role of the analysts has grown over time, as the program has matured.

The GBA analysts provided the initial input and the processes they used formed a baseline to determine what data should be collected. This period was followed by an

initial data collection which involved two analysts each working a long-term (9 months) task and producing one final product at the end. During this time, Glass Box data was distributed to the researchers monthly.

Knowingly, the initial data collection period was not realistic. It was designed to give researchers a feel for the data that Glass Box could collect and a chance to provide feedback to the developers. It was fully realized that analysts work on more than one task at a time, deal with competing priorities, unexpected interruptions, and produce more than one product, often simultaneously. For the period July-December 2004, there were four GBA analysts, one senior analyst with over 10 years experience and three junior analysts with less than 5 years experience, working in a multi-tasking environment (MTE).

The MTE collection effort began on 7 July 2004 and ended 10 December 2004. The period was divided into three 5-week cycles. Each analyst was given a focus area to monitor which spanned all three cycles. At the current time, the three cycles have been completed, and will be followed by a fourth cycle, which has already begun, and a fifth cycle.

The first two cycles involved two analysts each. These cycles were meant to allow procedures and protocols to be developed and adapted to meet the demands of the MTE. We learned during Cycle 1 that evaluators and researchers required strict protocols to enable them to match assigned tasks to the final intelligence reports. A simple task numbering system and file-naming convention were introduced in Cycle 2 that alleviated this problem.

Cycle 3 involved all four analysts and was designated as the baseline cycle. The analysts were unaided by any NIMD tools during this cycle. In future cycles, as new methods and tools are introduced, the analysts' performance can be compared to Cycle 3 data to determine the tools'/methods' effectiveness.

In order to make the environment as realistic as possible, the analysts were challenged. During the first three cycles (approx. 15 weeks) they received a total of 60 taskings and generated 72 intelligence reports. On any particular day, an analyst might have had only one on-going task or as many as five. Analysts were not aware when new tasks would be assigned. A number of lessons learned came from these cycles and are discussed in the next three paragraphs.

In anticipation of the MTE, the ability to switch tasks on-the-fly was added to the Glass Box instrumentation. It was believed that the analysts would be researching task 1, find information that related to task 2, and switch over to task 2 to deal with the new information. However, we discovered that analysts didn't actually work on more than one task at a time. The analysts would prioritize their tasks, normally giving the most weight to the task which was due first, and work only the highest priority task to completion. Only after that task was completed would they move on to the next task. There was very little switching between tasks, despite the number of simultaneous tasks the analysts were given or how much the tasks overlapped. If, during a longer term task, the analyst received a very short term task, the analyst would stop all work on the long term task and completely devote his/her time to the short task. This behavior seemed to be common to all four analysts.

After reviewing the MTE cycle data, evaluators at NIST had a number of suggestions for additional data capture. The majority of these recommendations would require more metadata about the tasks themselves be captured and stored. This metadata would include information on the type of report required by the task (e.g. bulletin, assessment, etc.), the projected amount of time to complete the task, the actual amount of time to complete the task, and feedback from the analyst's manager on the report and task execution.

Even as we digested the lessons of cycles 1-3, Cycle 4 began. The focus of Cycle 4 was not additional multi-tasking, but the introduction of structured analytic methods and their effect on the analytic process. During the first week of January 2005, the four analysts that participated in cycles 1-3 received formal training in the methods of evidence marshalling, structured argumentation and hypothesis generation. The first Cycle 4 task required the analysts to apply the structured methods to a task they had already performed during Cycle 1 or 2. The goals of this task were to determine if the analysts:

- Changed/modified their conclusions
- Found it more difficult to apply the methods to the real-world as opposed to classroom scenarios
- Absorbed the training and could apply it realistically

Glass Box data from Cycle 4 was not available at the time this paper was written. However, a debriefing of three of the analysts was conducted on January 13, 2005 after they had completed the first Cycle 4 task. A summary of this debriefing is provided below.

- Two analysts changed or modified their conclusions from their Cycle 1-2 analysis
- All analysts agreed that the method was easier to apply to classroom "event-based" scenarios then to real-world non-event-related tasks
- Additional training on assumption analysis as a step in the analytic process should be added to help analysts articulate key assumptions as required by the methodology
- Analysts needed to "question the question" to determine if the customer asked what was critical and germane to the task subject and form alternate hypotheses
- Decomposing hypotheses helped generate "new thinking" and find key indicators of activity
- The method helped analysts consider the forensics (social networks, relationships between activity and facilities, etc) of the problem
- The method consumed a lot of time and analysts wondered if it would useful for short term tasks

No changes to the Glass Box instrumentation have been identified yet from Cycle 4 activities. The NIST evaluators, however, have already expressed concern that they will not be able to pick out from the data instances of the analysts

applying the methodology since they had already observed the analysts using some elements of the methodology intuitively in their earlier tasks. However, until the Cycle 4 data is examined, it will be difficult to make recommendations to extend the instrumentation to better capture the use of structured methods.

Cycle 5 is currently scheduled to begin February 14, 2005. Cycle 5 will focus on collaboration, where more than one analyst is assigned to a task. It is expected that analyst-provided annotations will be very important to capture off-line collaborative meetings and discussions that cannot otherwise be captured by the Glass Box software. Additionally, we are creating a number of protocols that will aid in tracking multiple analysts working on one task:

- Each task will be given a common task number and multiple analysts assigned to it (rather than multiple tasks assigned to one analyst)
- Shared folders on the network will be established for each task to hold all generated research material and report drafts so all analysts can access them
- The use of instant messaging and chat rooms will be encouraged for discussions about the task so that Glass Box can capture these interactions
- An analyst work room will be established with a video camera to capture planning, brainstorming and other collaborations

The future after Cycle 5 has not yet been decided. It is expected that NIMD tools will be integrated into the Glass Box and analysts will begin using the tools in order to evaluate their effectiveness in the analytic process. Additionally, other analytic methodologies may be introduced and evaluated.

## 3. Glass Box Instrumentation

A continuing challenge has been to define requirements for automated data collection functions that are unobtrusive, yet robust and complete enough not only to capture lower-level data on human-computer transactions but also to shed light on the analyst's higher-level cognitive processes. Initially we focused on recording IA activities for a single analyst and a single task at a time and then evolved to more sophisticated capabilities for monitoring of multiple tasks simultaneously. Now we are moving toward data collection in collaborative IA environments.

The current Glass Box instrumentation captures analyst workstation activities including keyboard/mouse data, window events (active window, active application, location on screen, etc.), file open and save events, copy/paste events, and Web browser activity (URLs, page contents, images, queries, and query results). The Glass Box makes extensive use of a relational database to store time-stamped events and a hierarchical file store where files and the content of web pages are stored. The Glass Box "snatches"

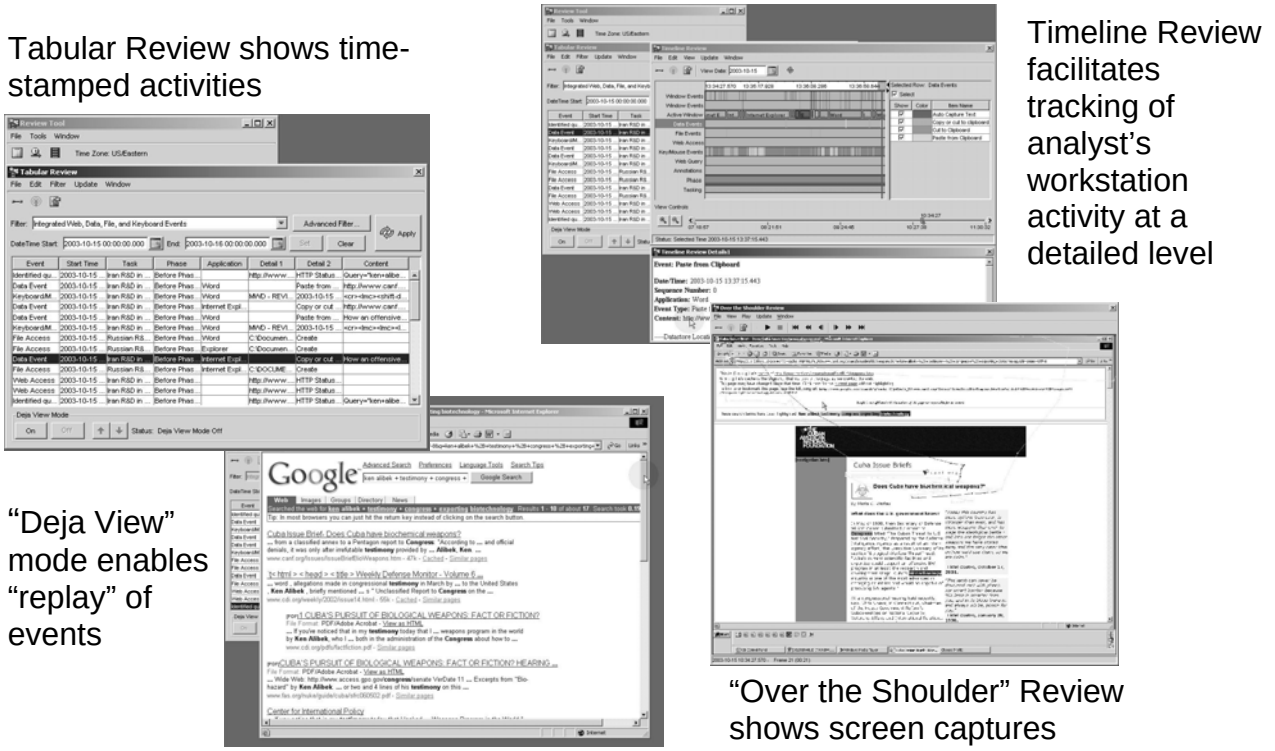


**Figure 1. Glass Box Review Tool**

a copy of every file the analyst opens and saves so we have a complete record of the evolution of documents. We also explicitly store the contents of every web page the analyst visits so we can recreate the content as it existed at the time the page was visited. We capture screen images to see what the analyst saw during the session. The Glass Box's Control Panel allows the analyst to suspend and resume recording to make sure the software does not inadvertently capture sensitive or proprietary information. Analysts also have the ability to delete data inadvertently recorded.

The instrumentation we have described to this point is relatively passive. Except for the responsibility of the analyst to turn on or turn off Glass Box recording, there is virtually no interaction with the analyst and no cooperation is required from the applications we have instrumented. Most of what this instrumentation captures is obtained by observing the application from the operating system level. However, on-line machine generated data does not reflect the entirety of analytic activity.

Much of the analytic process occurs off line and is important to modeling analysts and analytic processes. The Glass Box provides means for the analyst to record comments and notes about off-line activities. The analysts use the Glass Box annotation tool to note meetings, casual discussions, items of interest from the media, hypotheses being considered, plans for the day, etc. Analysts can also annotate and mark the relevance of citations, files, and excerpts from files. These annotations often provide the NIMD researchers with important context that supplements the recorded workstation activities, but they are written for processing by humans rather than by the NIMD applications.

The Glass Box also serves as a sensory mechanism for the NIMD applications through our Application Programming Interface (API). The API provides applications with access to user interactions and system activity, allows applications to log activities of interest, and allows applications to communicate through our publish/subscribe mechanism. This allows NIMD applications to find out what the analyst is doing in near real time, determine a course of action, and to coordinate their response. Researchers and those evaluating the applications can use the resulting data to determine how the application actually got used, where the analysts had problems, where they took unexpected paths, what the conditions were under which problems developed, etc.

## 4. Distilling Analytic Events

The Glass Box collects low-level data that in and of itself has limited value towards studying the analysis process. This data becomes interesting, however, when it is aggregated, distilled or combined to indicate analytic events. We define an analytic event as a primitive or composite event type that is significant to the analytic process. The NIMD community is still wrestling with what data may be defined as an analytic event. Some data has immediate and obvious significance, such as issuance of a search engine query. Other data, like individual keystrokes, are only interesting after some number of them is combined into meaningful structures such as words, sentences, or paragraphs. For example, researchers at the National Institute of Standards and Technology (NIST) studied the growth of an intelligence report over time by combining individual time stamped characters and cut/paste events captured in the Glass Box to view the differences in how analysts generate intelligence reports (Scholtz *et al.* 2004). The challenge is to define at which level of aggregation these low-level events become analytic events. For instance, many researchers have expressed interest in knowing when an analyst finds a particular document relevant to the current task. There are a number of indicators of relevance: saving the document, printing the document, cutting and pasting from the document into a report, and reading the document (as opposed to scanning the document). Glass Box does not currently capture a "reading" event, but it may be possible to extrapolate a reading event. Sanquist, *et al*. (2004) showed that activities could be characterized by their event dwell times, and specifically showed that by combining the time an analyst has a document displayed with other activities, one could postulate that the analyst is actually reading the displayed document.

Although specific analytic events have not yet been identified, the GBA team has begun to develop event types to help evaluators, developers and researchers characterize entries logged to the Glass Box. Each event may have zero to many types:

- Functional (e.g., copy, paste, hide, show, open, close, etc.)
- User-related (e.g., dialogue with the user, reporting an error condition to the user, annotations made by user, read/view document, etc.)
- Errors (e.g., bugs, getting to places the user never should have gotten to, unanticipated conditions, memory problems, etc.)
- Query-related (e.g., perform query, query results, links returned, links followed, etc.)
- Agent-related (e.g., assumptions, parameters passed, etc.)
- User-model related (e.g., behavior, interest, bias, etc.)
- Metrics-related (e.g., data used to derive solution time, number of queries, number of documents read, comparison to expert, rate of information growth, accuracy, quality (including confidence in recommendation, etc.))
- Time-related (start, stop, iterations performed, etc.)
- Application-specific (e.g., cell operations, sandbox operations, etc.)

Some events are common to many applications (e.g., copy, cut, paste, start, end, hide, show, query, annotation, error message, etc.) Other events will be specific to a single application (e.g., create object, add item to cell, agent assumptions, etc.). Using common event types across applications allows applications to be compared (e.g., how many documents did the analyst actually read using Application A compared to Application B?).

To better support access to analytic events, the Glass Box development team is migrating to a Services Oriented Architecture that uses web services and an Analytic Event Server (AES) to store and disseminate analytic events to researcher tools across firewall boundaries. Low-level events will be collected and aggregated on the workstation and then passed to the AES for persistent storage. Through web services messaging, NIMD tools can request and receive notification in near real-time when selected analytic events occur. The tools may also log their own tool-generated events to the AES for sharing with other NIMD applications. Furthermore, since the analytic events are time stamped and permanently stored, a tool that has been off-line can be "caught up" with activity that occurred since the tool was last on-line.

## 5. Storage and Retrieval

Throughout this project, the Glass Box team's goal has been to provide a rich and meaningful set of data that could be used for a variety of purposes, including some that were not anticipated. We complement this with software to provide quick and easy access to a variety of views of the collected data.

The Glass Box Review Tool, as shown in in Figure 1, provides analysts and researchers with a quick review and analysis capability, including the capability to observe the progress of activity over time, repeated patterns over time, information clusters (i.e., areas where analysts repeatedly spend time), breadth of scanning, and narrowing of focus.

The Review Tool provides different and complementary ways of visualizing the Glass Box data. The **Tabular Review** window displays a chronological record of recorded events that can be explored in detail. The display can be filtered in various ways (using specialized queries) to study different types of activities and relationships in the data. (These queries are also available through our API to allow applications access to the same data.) The Tabular Review window also allows the events to be displayed in a context and view that is very close to what was seen by the analyst (we call this view "deja view"). **Over-the-Shoulder** and **Timeline Review** windows provide additional ways of visualizing the data that help to develop an understanding of the structure, timing, and content of analyst activity. Over-the-Shoulder Review provides screen images of what the analyst was seeing on his or her workstation at the selected point in time (currently one image per second). The Timeline Review provides a time-based visualization of multiple events in relation to each other. This view shows the flow of information across activities, the specific points in time where analysts performed mouse actions and keyboard inputs, and the resulting effects on applications and windows.

## 6. Use of Glass Box Data

The NIMD researchers are responsible for exploring the Glass Box data, making discoveries within the data, and applying the data to understand the higher cognitive processes involved in intelligence analysis. Several advances amongst the NIMD researchers have resulted. Examples are highlighted in this section.

**NIST** uses the Glass Box in both its NIMD and non-NIMD work. For its NIMD work, Glass Box serves as a "surrogate observer" while conducting evaluation experiments on NIMD applications. The Glass Box logs workstation activities to a level of detail that no human observer could achieve and provides the ability for the human observers to review those observations at many different levels of detail. NIST's evaluation of the Glass Box found that it does a good job in capturing workstation activities (and is much better than video) but does not capture sufficient data about off-line activities. NIST is encouraging NIMD researchers to have their applications capture the same kind of data that the Glass Box currently provides on activities within an application. Through use of API, Glass Box can provide valuable insight to the researchers and evaluators on usability and other problems.

The **Palo Alto Research Center (PARC)** created printable/scrollable visualizations of the analytic processes inherent in the Glass Box data. PARC is working on software to create a "personalized search profile" for each analyst from the Glass Box data. PARC has also used the Glass Box data to examine window thrashing (rapid switching between open windows) and recommended a seamless, multi-monitor system to improve the display of data to the analyst (Mackinlay and Royer, 2004).

**Sarnoff** generated "search sagas" using Glass Box-recorded queries, query results, and the pages visited by the analyst. These sagas are used to evaluate analyst interest in order to predict the next relevant document. Sarnoff also is looking into how to use Glass Box data to evaluate their anti-bias mechanism aimed at helping an analyst find more relevant documents.

**Cycorp** has integrated the Glass Box schema and data into its Cyc knowledge base. Cycorp implemented a prototype application to perform a shallow parse on some of the Glass Box text data, identify corresponding Cyc concepts, and use these as indicators of the analyst's information needs and what actions the application should take to support the analysis process.

**Oculus** evaluated their tool's performance compared to unaided Google™ searches captured in the Glass Box.

Oculus is also involved in an integration experiment to be conducted in 2005 with Altarum and Sarnoff, to demonstrate communication through the Glass Box API.

**Computer Science Innovations, Inc. (CSI)** focused on concept extraction and modeling based on context of current analyst activity using Glass Box captured analyst annotations, cut/paste operations, queries and documents retrieved, and used time stamp data to calculate concept (event) density with the goal of minimizing the path between concepts. A key discovery has shown the analyst in multiple analytic phases simultaneously (i.e., planning, collecting, reporting, etc.).

Our Glass Box API is already being used by NIMD researchers. Oculus has successfully used our logging capability. Georgia Tech Research Institute has integrated their case-based reasoning tool with the Glass Box to store data on the uses of particular cases and to store run-time statistics for performance evaluation and analysis.

## 7. Conclusions

We have learned much in the process of instrumenting the analysis process. What we capture is representative of the on-line activities performed by analysts. Through our review tools and specialized queries, users can access the diverse data that the process generates and make sense of it. Through our annotation capability, analysts can document off-line activity. Through our API, applications can communicate between themselves, find out what analysts are doing, and determine courses of action to aid the analyst. The Glass Box and its data can also be used to evaluate how applications are performing and documenting how users are actually using the applications.

The GBA analysts and their work in the Glass Box continue to provide new insights into the analytic process. The GBA environment has matured from a single-task, single-analyst scenario to a full-fledged multi-tasking environment, where the focus is on providing a realistic experimental environment that simulates the real-world open source analytic process used in the Intelligence Community (IC). We have further introduced structured analytic techniques such as evidence marshalling, structured argumentation and hypothesis generation, and have plans to examine analytic collaboration. The data gathered from these studies has led researchers and evaluators to discover new information about how analysts work. We saw that the analysts generally focused on only one task at a time despite a high pressure multi-tasking environment. Evaluators were also able to detect that the analysts intuitively followed some elements of the structured analysis methodologies before they were formally trained in them. The GBA data has allowed other researchers to examine windows thrashing, visualization of the analytic process, report generation/growth over time, analytic bias, and document relevancy, to name a few.

The objective of the GBA experiment is to enable the study of the analysis process, rather than studying the differences between individual analysts so our sample size is relatively small. Once information on the process is captured, NIST uses the Glass Box data to compare how multiple analysts approach the same problem by repeating the GBA experiments (analytic tasks) using analysts from the IC. In this and the other ways discussed in this paper, GBA data reveals the analysis process.

The full richness of the Glass Box data and its information on the analytic processes has not yet fully been exploited. For example, we suspect that tacit and prior knowledge and hypothesis generation can be derived from Glass Box data. We also believe that the data has value for the kinds of studies discussed in (Mackinlay and Royer, 2004).

Others in the IC have expressed interest in the Glass Box and several have the software and regularly receive our data releases. In fact, the non-NIMD user community is growing steadily. Others in the IC interested in obtaining the Glass Box software or data should contact the authors for more information on how to obtain the Glass Box.

## Acknowledgements

This work was funded by the Advanced Research and Development Agency's Novel Intelligence from Massive Data Program.

## References

Cowley, P. J., Nowell, L. T., Scholtz, J. 2005. Glass Box: An Instrumented Infrastructure for Supporting Human Interaction with Information. Big Island, Hawaii: 3-6 January 2005.

Gorton, I. and Haack, J. N. 2004. Architecting in the Face of Uncertainty: An Experience Report. 26th International Conference on Software Engineering. Edinburgh, Scotland: 23-28 May 2004, 543-551.

Mackinlay, J. D. and Royer, C. August 2004. Log-based Longitudinal Study Finds Window Thrashing. Technical Report UIR-2004-06, Palo Alto Research Center.

Sanquist, T. F., Greitzer, F. L., Slavich, A., Littlefield, R., Littlefield, J., and Cowley, P. (2004). Cognitive Tasks in Intelligence Analysis: Use of Event Dwell Time to Characterize Component Activities. Human Factors and Ergonomics Society 48th Annual Meeting. New Orleans, LA: 20-24 September 2004.

Scholtz, J., Morse, E., and Hewett, T. 2004. In Depth Observational Studies of Professional Intelligence Analysts. Human Performance, Situation Awareness and Automation Technology Conference. Daytona Beach, FL: 22-25 March 2004.