Ordinal Structure of the Generic Analytic Workflow: A Survey of Intelligence Analysts

Mandeep K. Dhami
Department of Psychology
Middlesex University
London, UK
m.dhami@mdx.ac.uk

Kathryn E. Careless
Department of Psychology
Middlesex University
London, UK

***Abstract*—We present a model of the generic analytic workflow and measure the extent of ordinal structure that analysts apply to analytic work, as well as how this is affected by their training and experiences. The workflow comprises six stages that follow from one another: capture requirements, plan analytic response, obtain data, process data, interpret outputs, and communicate conclusions. A survey of 144 intelligence analysts revealed that only 16% structured their workflows in a logical ordinal way. The extent of ordinal structure applied to analytic work was unrelated to analytic thinking training, years of analytic experience, and proportion of time spent working collaboratively. These findings have implications for the training and assessment of analysts, as well as for the design of analytic tools.**

***Keywords*—Intelligence analysis; analytic workflow; sense-making**

I. Introduction

In the present paper, we propose a model of the generic analytic workflow, and shed light on the extent of ordinal structure that analysts apply to analytic work, as well as how this may be affected by their training and experiences. Below, we first review literature on the analytic workflow and present our model. We then report a study of how analysts structure their workflows. Finally, we discuss the implications of these findings, and suggest directions for future research.

II. The Analytic workflow

A. Past Research on Analytic Workflows

Researchers have aimed to identify stages of the analytic workflow, primarily to inform the design of analytic technology [1][2][3][4][5][6][7][8][9]. Phillips et al. (2001)[7] used a literature review and interviews with an unspecified number and type of participants to develop a seven phase model: define problem, identify knowledge base, target location of information, select intelligence mode (e.g., passive), collect information, analysis of competing hypotheses [10], and report. The model includes 'mini-phases' representing a feedback loop. The main limitations of this study are that it was based on an unspecified sample and it excludes analysts' processing of data.

Elm et al. (2005)[2] conceptualized the intelligence analysis process as three primary iterative cognitive functions. Down-collect involves collecting data. Conflict and collaboration generate interpretations of the data. Hypothesis exploration constructs hypotheses to help interpret or explain the findings. However, this model does not specify when analysts capture requirements, plan their analytic response, and communicate their conclusions. Indeed, others have similarly argued for the importance of including a capture requirements stage (called problem formulation or information needs management; [3][9]). Farry et al. (2011)[3] also recommends including an end stage called decision selection which would be part of communicating conclusions.

Pirolli and Card (2005)[8] interviewed two experienced intelligence analysts using a cognitive task analysis. The analysts also completed a simulated analytic problem while following think-aloud protocols. The findings led to a 16-step model as follows: external data sources, search and filter, search for information, 'shoebox', read and extract, search for relations, evidence file, schematize, search for evidence, schema, build case, search for support, hypotheses, tell story, re-evaluate, and presentation. Steps 1 to 7 are included in a foraging loop and steps 8 to 16 are in a sense-making loop. Bier et al. (2008)[1] extended the model to collaborative analysis by sharing and viewing of others' work. Unfortunately, Pirolli and Card's model was based on only two analysts. It also focuses almost exclusively on analysts obtaining data, processing it and interpreting the outputs of such processing, thus excluding other crucial stages (i.e., capturing requirements, planning the analytic response and communicating conclusions).

Klein, Phillips, Rall, and Peluso (2007)[11] proposed what they call the data/frame model of sense-making that is involved in the intelligence analysis process (see also [5]). According to this, analysts construct a frame (story or mental script) from available data and existing frames. Sense-making attempts to connect data to a frame; elaborate the frame; question the frame; preserve the frame; seek a frame; re-frame; and compare multiple frames. Sense-making activities involve discarding existing data; gathering/inferring new data; judging the plausibility and quality of data; distorting data or re-interpreting it; inferring and revising relationships between data; detecting and tracking anomalies/inconsistencies in data; and recovering discarded data. This model, however, excludes analysts planning their analytic response and communicating their conclusions. It also confounds obtaining data, processing it, and interpreting the outputs.

Moore (2011)[6] suggested that intelligence sense-making involves at least five overlapping activities: planning; foraging

The present research was funded by HM Government.

for information; marshalling, which involves processing data; understanding, and communicating. This however confounds analysts capturing requirements with their planning their analytic response.

Finally, Kang and Stasko (2011)[4] conducted an observational study and semi-structured interviews with three teams of intelligence studies students (N = 14); each performing different hypothetical analytic tasks. This revealed four phases: building an explicit conceptual model of the problem; collecting data; analysis of the data; and producing the report/presenting findings. The main limitations of this study are that it included a small sample of non-analysts, and the model is based on collaborative analysis. It also excludes analysts planning their analytic response, and confounds analysts processing data with them interpreting the outputs.

B. *The Generic Analytic Workflow*

In order to overcome some of the limitations of the above work, we developed a model of the *generic* analytic workflow. It is generic because it applies to various types of analysis (e.g., HUMINT, SIGINT, multi-source), conducted individually or in teams, for various purposes (e.g., strategic, tactical). The workflow is separated into *six* different stages of activity that follow one another, and which represent the full spectrum of analytic work. The six stages are: capture requirements, plan analytic response, obtain data, process data, interpret outputs, and communicate conclusions (see Figure 1).

The *capture requirements* stage is about understanding the customer's viewpoint, what outcome the customer wants to achieve, and challenging this if necessary. The *plan analytic response* stage is about identifying the analytic lines/hypotheses, the methods for evaluating these, and how effective and efficient they may be, as well as prioritizing how to proceed. The *obtain data* stage is about extracting and selecting the relevant data from the most appropriate sources in the most efficient manner, as well as establishing new sources of data if necessary. The *process data* stage is about manipulating the data using relevant analytic tools and techniques, including reformatting it. The *interpret outputs* stage is about evaluating alternative explanations for the data, constructing a logical argument to support the conclusion(s) drawn, determining the degree of confidence in these conclusions, and identifying any ambiguities. Finally, the *communicate conclusions* stage is about communicating the outcome of analysis in a clear and meaningful format, distinguishing fact from inference, and expressing uncertainty and confidence.

We confirmed the existence of the six stages by referring to the past research reviewed above, as well as documented best practice and performance standards from the intelligence community [see e.g., [12][13]]. In addition, as part of another project, we confirmed the description of each stage via discussions with a small sample of experienced analysts, managers of analysts, and trainers in analytic methods [14].

We propose that the six stages follow from one another. Depending on the scale of the analytic problem and the analysts' experience, analysts might pass through some of the stages very quickly, while remaining at other stages for some time. The workflow might be linear for simple, discrete analytic problems. Indeed, it is logical to expect that an analyst would begin by fully and precisely capturing a customer's requirements. He/she might then be expected to carefully plan an analytic response that involves identifying and obtaining the necessary data, before processing it using relevant techniques and tools. Finally, an analyst might be expected to interpret the output of such analyses before communicating the conclusions to a decision-maker appropriately.

However, the workflow might also be iterative for more complex analytic problems, or set of inter-connected problems making the task dynamic in nature [15]. For instance, in Figure 1 an analyst might realize that the data goes beyond that necessary to fulfil the customer's requirements, and so 'loops back' to the first stage in order to further clarify and/or expand these requirements. This will then mean that the analyst re-plans the analytic response and obtains further data before re-processing. Others have also noted the importance of iteration and non-linearity of the analytic workflow (see [2][4][6]), and explicitly included feedback loops in their models [5][7][8].

The important point is that analysts should not skip or omit a stage. This is because a lack of ordinal structure may result in inconsistent working practices as well as analytic work that lacks transparency, thus difficult to review or audit. In addition, divergence from an ordinal structure can lead to analytic products that do not fully satisfy the requirements i.e., are ineffective. Finally, a lack of ordinal structure may lead to inefficient and resource-intensive (i.e., costly) working practices.

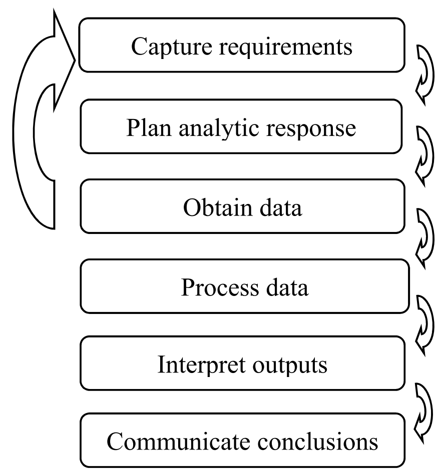

Fig. 1. A model of the generic analytic workflow (also depicting an example of a 'loop back')

III. THE PRESENT STUDY

We examined the ordinal structure that analysts apply to analytic work. We also explored the relationship between how analysts' ordered the workflow and their analytic thinking training, years of experience, and proportion of time spent working collaboratively. To-date, no-one has examined the first issue. Regarding the second issue, one might expect analysts' training and experiences to be associated with how they perform analytic tasks.

IV. METHOD

A. Participants

Participants were 144 UK intelligence analysts. Around half (56.8%) of the sample were male.[1] The average age was 37.27 (*SD* = 10.28). Nearly all (92.8%) of the sample was employed to work full-time. Thirty-three percent had undertaken some training on analytic thinking. On average, the sample had 5.18 (*SD* = 3.38) years of experience working in the intelligence community. The sample spent on average 35.68% (*SD* = 22.31) of their time per week working collaboratively/as part of a team.

B. Survey

A structured survey was designed for the present study by a team of experienced analysts (including managers) and analytic trainers in collaboration with the authors. The survey comprised four parts. The first and last parts of are relevant to the present study. (The second and third parts measured how often analysts applied specific analytic strategies to solving tasks along the analytic workflow and measured their thinking style, respectively; the findings are reported in [16]).

As mentioned, one part of the survey captured the extent of ordinal structure that analysts applied to analytic work. It listed hypothetical activities that represented the six stages of the workflow (i.e., capture requirements, plan analytic response, obtain data, process data, interpret outputs, and communicate conclusions). An early version of the survey was pilot tested on a sample of 60 analysts (they did not participate in the actual study). This revealed that the analytic scenario and activities were generally representative of the six main stages of the workflow. Analysts in the present study were asked to rank the activities in the order they would usually carry them out in response to a brief analytic scenario.

Another part of the survey captured information about the analysts on a range of variables i.e., their gender, age, work status, analytic training courses undertaken, years of experience working in the intelligence community, and proportion of time spent per week working collaboratively/as part of a team ('collaboration' was not explicitly defined).

C. Procedure

The survey was available online for a two-week period on the organization's intranet. Analysts were recruited on a voluntary basis via advertisements on the intranet of an intelligence organization. They were expected to complete the survey during their workday. This took approximately 25 minutes. Participation in the study was anonymous.

V. FINDINGS

A. Ordinal Structure of Analytic Work

The main aim of the study was to determine the extent of ordinal structure that analysts apply to analytic work. Sixteen percent of analysts ranked all of the six activities presented in the survey in the logical order (i.e., capture requirements, plan analytic response, obtain data, process data, interpret outputs, and communicate conclusions). Seven percent said they would perform the activities in the first half of the workflow (i.e., capture requirements, plan analytic response, and obtain data) in the logical order. Another seven percent said they would perform the activities in the second half of the workflow (i.e., process data, interpret outputs, and communicate conclusions) in the logical order. The vast majority (70.1%) of analysts, however, applied less ordinal structure to analytic work.

Further analysis of the responses of these 70.1% of analysts revealed several different ways in which they ordered the activities. In particular, analysts said they would obtain data after capturing requirements or would plan the analytic response after processing the data. Analysts also said they would interpret outputs after obtaining data or would interpret the outputs directly after planning the analytic response. Therefore, analysts who did not structure their workflows in a logical ordinal fashion tended to delay planning and/or start interpretation prematurely.

B. Analysts' Training and Experiences and Ordinal Structure of Analytic Work

A secondary goal of the present study was to explore the relationship between how analysts' order the generic analytic workflow and their analytic thinking training, years of work experience, and proportion of time spent working collaboratively. Whether or not analysts had undertaken an analytic thinking course did not have a statistically significant effect on the extent of ordinal structure they applied to analytic work, $p > .05$. Similarly, years of experience working in the intelligence community, and the proportion of time spent per week working collaboratively were not significantly associated with the extent of ordinal structure applied, $ps > .05$.

VI. DISCUSSION

In the present paper, we have proposed an evidence-based model of the generic analytic workflow that covers the full spectrum of analytic work. We found that when presented with a representative analytic scenario which involved six activities, one pertaining to each of the six stages of the workflow, only a minority of analysts structured analytic work in a logical ordinal fashion. It is likely that fewer analysts would apply a logical ordinal structure to analytic work, and that the extent of structure would be less, for more complex problems. Future research ought to examine the effect of task complexity on ordinal structure of analytic work. Complexity might be variously defined, for instance, in terms of information overload, time pressure, and choice options (e.g., choice of data sources or data processing tools) [17].

The present findings identified four key areas of concern in relation to how analysts may structure their workflows. First, some analysts only managed to structure the first or second half of the workflow in a logical order. This suggests that analysts might start an analytic task with some ordinal structure, but then 'lose the thread' of the workflow or that they might start analytic work in a less orderly fashion, but eventually 'grasp the thread' of the workflow. Further research needs to be conducted to identify the factors that might stimulate this workflow breakdown and coherence, respectively. For instance, a breakdown might occur if

[1] Full demographic data was only available for 113 participants.

analysts are unfamiliar with analytic tools and techniques, whereas a workflow might become more structured as analysts start to understand the data.

Second, analysts were prone to 'skipping' or delaying the planning analytic response stage. Some moved directly from capturing requirements to obtaining data, while others planned after they had processed the data. This suggests that analysts may be keen to 'jump in' and collect and process data before planning their work. They might then use the data and output of the processing to inform what they ought to do next. This is problematic because if plans are being driven by the data an analyst obtained from an intuitive or even random search, the whole analytic process is skewed by the initial choice of data sources, queries used, and the data obtained, rather than a thoughtful consideration of the analytic question. In addition, planning the analytic response after processing data could lead to biases in, for example, obtaining further data.

Third, analysts in the present study were also prone to prematurely interpreting outputs before actually processing the data. This suggests that analysts may rely on intuitive responses to the analytic problem, rather than on critical thinking. Such an approach may mean that analysts resort to drawing (familiar) conclusions to familiar problems, rather than considering unfamiliar conclusions or directly fulfilling a customer's requirements.

Finally, the fact that the extent of ordinal structure that analysts applied to analytic work was unrelated to their analytic thinking training, years of work experience in the intelligence community, and proportion of time spent per week working collaboratively, is of concern. To some extent this might be explained by the intelligence community's assertion that intelligence analysis is not a science [18].

From a practical standpoint, analytic training could emphasize the importance of applying an ordinal structure to analytic work (e.g., transparency). Assessments for career progression could also measure competency in structuring analytic workflows. Leaders of analytic teams could attempt to impose greater structure on collaborative analytic activities. Finally, analytic tools designed to support and aid analysts in performing certain activities such as obtaining data and processing it, could also help guide them through their workflows in a logical fashion.

Although the present study has focused on the importance of applying an ordinal structure to the analytic workflow, it is important to recognize that workflows for complex or dynamic problems may require iteration. Future research can explore how analysts iterate through analytic problems and what factors initiate loop backs (see e.g., [15]).

Although analysts who skip or omit a stage, even when iterating or looping back, may work faster, their analytic outputs might be less effective. Future research could explore the effects of a lack of ordinal structure. For instance, to what extent does a lack of ordinal structure affect the quality of the analytic output? How does a lack of ordinal structure affect the transparency of analytic processes? Is a lack of ordinal structure associated with inefficient working practices? The findings of such research can have implications for the development of best practice standards.

Acknowledgment

Thank you to Ian Belton for his research assistance.

References

[1] Bier, E. A., Card, S. K., & Bodnar (2008). Entity-based collaboration tools for intelligence analysis. In D. Ebert, & T. Ertl (Eds.), *IEEE Symposium on Visual Analytics Science and Technology (VAST) 2008* (pp. 99-106).

[2] Elm, W., Potter, S., Tittle, J., Woods, D., Grossman, J., & Patterson, E. (2005). Finding decision support requirements for effective intelligence analysis tools. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 49*, 297-301.

[3] Farry, M., Carlson, E., & Mahoney, S. (2011). The mutual support function model: A cognitive model for intelligence analysis supporting irregular warfare. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 55,* 880-884.

[4] Kang, Y. & Stasko, J. (2011). Characterizing the intelligence analysis process: Informing visual analytics design through a longitudinal field study. In G. Santucci & M. Ward (Eds.), *IEEE Conference on Visual Analytics Science and Technology (VAST) 2011* (pp. 21-30).

[5] Klein, G., Moon, B., & Hoffman, R. R. (2006). Making sense of sensemaking 2: A macrocognitive model. *IEEE, Intelligent Systems, 21*, 88-92.

[6] Moore, D. T. (2011). *Sensemaking: A structure for an intelligence revolution*. Retrieved from http://www.nil-u.edu/ni_press/pdf/Sensemaking.pdf

[7] Phillips, J., Liebowitz, J., & Kisiel, K. (2001). Modeling the intelligence analysis process for intelligent user agent development. *Research and Practice in Human Resource Management, 9*, 59-73.

[8] Pirolli, P., & Card, S. K. (2005). The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. *Proceedings of International Conference on Intelligence Analysis, 5*, 2-4. Retrieved from http://vadl.cc.gatech.edu/documents/2_card-sensemaking.pdf

[9] Roth, E. M., Pfautz, J. D., Mahoney, S. M., Powell, G. M., Carlson, E. C., Guarino, S. L., Fichtl, T. C., & Potter, S. S. (2010). Framing and contextualizing information requests: Problem formation as part of the intelligence analysis process. *Journal of Cognitive Engineering and Decision Making, 4,* 210-239.

[10] Heuer, R. (1999). *Psychology of intelligence analysis*. Washington, DC: US Government Printing Office.

[11] Klein, G., Phillips, J. K., Rall, E. L., & Peluso, D. A. (2007). A data-frame theory of sensemaking. In R. R. Hoffman (Ed.), *Expertise out of context: Proceedings of the 6th International Conference on Naturalistic Decision Making* (pp. 113-155). Lawrence Erlbaum Associates.

[12] UK MOD (2013). *Quick wins for busy analysts*. Published by Defence Intelligence, UK.

[13] US Government (2009). *A tradecraft primer: Structured analytic techniques for improving intelligence analysis.* Retrieved from http://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/Tradecraft%20Primer-apr09.pdf

[14] Dhami, M. K., & Careless, K. (2015). *Development and validation of the analysis support guide*. Manuscript in preparation.

[15] Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica, 81*, 211–241.

[16] Dhami, M. K., & Careless, K. E. (2015). *Strategies for Solving Analytic Tasks: A Survey of Intelligence Analysts.* Manuscript submitted for publication.

[17] Wood, R. E. (1986). Task complexity: Definition of the construct. *Organizational Behavior and Human Decision Processes, 37*, 60-82.

[18] Dhami, M. K., Mandel, D. R., Mellers, B. A., & Tetlock, P. E. (in press). Improving intelligence analysis with decision science. *Perspectives on Psycholological Science.*