# A cognitive prosthesis for complex decision-making

Sébastien Tremblay[1], Jean-François Gagnon[2], Daniel Lafond[2], Helen M. Hodgetts[1, 3], Maxime Doiron[1], & Patrick Jeuniaux[1]

[1]École de psychologie, Université Laval, Québec, Canada

[2]Thales Research & Technology Canada, Québec City, Québec, Canada

[3]Department of Applied Psychology, Cardiff Metropolitan University, Cardiff, UK

Word count (excluding references, abstract, acknowledgements, figures and tables) : 5 676

Correspondence to:

Sébastien Tremblay, PhD, École de psychologie, Université Laval, 2325 rue des Bibliothèques, Québec (QC), G1V 0A6, Canada. E-mail: sebastien.tremblay@psy.ulaval.ca. Phone: 418-656-2131 x2886; Fax: 418-656-3646.

Abstract

While simple heuristics can be ecologically rational and effective in naturalistic decision making contexts, complex situations require analytical decision making strategies, hypothesis-testing and learning. Sub-optimal decision strategies – using simplified as opposed to analytic decision rules – have been reported in domains such as healthcare, military operational planning, and government policy making. We investigate the potential of a computational toolkit called “IMAGE” to improve decision-making by developing structural knowledge and increasing understanding of complex situations. IMAGE is tested within the context of a complex military convoy management task through (a) interactive simulations, and (b) visualization and knowledge representation capabilities. We assess the usefulness of two versions of IMAGE (desktop and immersive) compared to a baseline. Results suggest that the prosthesis helped analysts in making better decisions, but failed to increase their structural knowledge about the situation once the cognitive prosthesis is removed.


*Keywords:* Complex decision making; Visual analytics; Knowledge representation

## 1. Introduction

It has been demonstrated on numerous occasions both in field studies and laboratory experiments that human decision-makers confronted with complex situations fail to perform satisfactorily despite their well-intended efforts (Frensch and Funke, 1995; Gonzalez, Vanyukov, and Martin, 2005; Osman, 2010a; Quesada, Kintsch, and Gomez, 2005; Yasarcan, 2009). Using computer-simulations of complex situations, Dörner (1996) noted particular examples of behaviors leading to successful performance (e.g., active learning and hypothesis testing) as well as numerous instances of poor behaviors leading to failure, which were generally linked to cognitive limitations and poor understanding and/or decision-making strategies (e.g., thinking in terms of isolated cause-and-effect relationships). Analysts and decision-makers confronted with complex situations could thus benefit from external support tools to help overcome cognitive limitations and facilitate broader situational understanding (e.g., interactive relationships, and projection of future consequences). The current study assesses whether IMAGE – one such cognitive prosthesis based on visual analytics – can augment analysts' structural knowledge and encourage optimal decision-making strategies.

Complex systems are characterized by uncertainty and non-linear interactions (Blech and Funke, 2005; Diehl and Sterman, 1995; Forrester, 1993) making it difficult to understand relations between elements. Furthermore, consequences of actions are often delayed in time and diluted by natural dynamic changes (Karakul and Qudrat-Ullah, 2008), while feedback may also be distorted, subject to misinterpretation, or imperceptible (Sterman, 2006). The dynamics of such systems are determined by their underlying structure, so knowledge about the causal relations between the elements comprising such systems – referred to as *structural knowledge* (Davis, Curtis, and Tschetter, 2003) – is critical for performing effective decision-making in this context (e.g., Blech and Funke, 2005; Gagnon et al., 2012). However, decision-making is not solely dependent on the quality of structural knowledge, as heuristics

– simple but effective strategies that do not require a profound knowledge of a situation – may also be used and can often lead to good performance in certain task ecologies that favour the simplicity principle (Gigerenzer and Gaissmaier, 2011; Gigerenzer and Selten, 2001; Shah and Oppenheimer, 2008). Of course, not all task ecologies favour simple strategies. For example, many contemporary organizations are complex sociotechnical systems that require analytically-derived management guidelines (Righi and Saurin, 2015). While simple heuristics can be ecologically rational and effective in various naturalistic decision making contexts, it follows from Ashby's law of requisite variety (Ashby, 1968) that complex situations require analytical decision making strategies, hypothesis-testing and learning.

Heuristics may introduce biases that can lead to substandard decisions and failure of strategic decision-making in complex situations such as health care (e.g., Agyepong et al., 2012), military strategic decision-making (e.g., Cohen, 2012), foreign policy making (e.g., Mitchell and Massoud, 2009) and macro-economy (Stekler, 2007). When dealing with such complex systems, heuristics fail to integrate sufficient complexity, and often generate less than satisficing outcomes (Betsch, Fielder, and Brinkmann, 1998; Betsch et al., 2001; Betsch et al. 2004). The limitations associated with intuitive heuristics have been referred to as cognitive "pathologies" (Cooper, 2005; Heuer 1999), with "pathological" behaviors including excessively reactive decision-making (focusing on fixing salient problems, i.e., a *firefighting* approach), lack of hypothesis testing, failure to consider potential side-effects or long-term effects of decisions, focusing on the present situation rather than on developmental trends, linearly projecting the situation into the future, searching for unique "one-factor" causes to problems, thematic vagabonding (focusing successively on different sub-problems with no coherent plan), and encystment (focusing on a single sub-problem) (Dörner, 1996). Decision-making quality may be helped by external tools that can support the development of structural knowledge rather than the use of heuristics.

### *1.1.Cognitive prostheses*

Cognitive prostheses are tools designed to augment cognition by offloading part of the information processing or representation requirement onto external artifacts (Cooper, 2005; Heuer, 1999). External cognition refers to the use of (mainly visual) representations to (1) reduce cognitive effort (computational offloading), (2) make problem-solving easier by *re-representing* information in a more tractable form, and (3) guide inferential reasoning about the underlying situation using graphs (see Scaife and Rogers, 1996). Tools supporting external cognition may promote the use of more analytical reasoning techniques over simple heuristics (Arias-Hernandez, Green, and Fisher, 2012), or help overcome cognitive bounds such as data overload and confirmation bias (Heuer, 1999; Johnston, 2005).

### *1.2. IMAGE – A cognitive prosthesis*

The IMAGE system (Lizotte et al., 2012) – so named to reflect its emphasis on visual representation – is a set of advanced visual analytics technologies to help improve analysts' understanding of complex situations by fostering the use of analytical reasoning strategies. IMAGE provides the user with added computational resources designed to support the adoption of "stronger" analytical methods of reasoning as opposed to "weaker" intuitive methods (see Bryant, Webb, and McCann, 2003). In order to achieve this goal, IMAGE provides three functions: (1) interactive simulations for hypothesis testing, (2) enhanced visualizations and (3) knowledge representation. Together, these functions allow the user to experiment with a simulation model to better understand a complex situation's dynamics. The user can manipulate the situation parameters and potential decisions in different simulation runs to observe the different outcomes. The user then attempts to discover trends, tipping points, and trade-offs using the interactive visualizations. Finally, the user captures his insights and his understanding of the complex situation in the knowledge representation component. This knowledge discovery process is not expected to operate in a linear sequential fashion, but rather as a series of iterations going back and forth across these different components.

### *1.2.1. Interactive simulation*

When acquiring structural knowledge, “direct” learning involving active interaction with the environment may be more effective in complex settings than (vicarious) learning by observing the interventions of others – i.e., indirect learning about the environment (Lagnado and Sloman, 2004; Osman, 2010b). Indeed, cognitive studies examining causal learning processes suggest that structural knowledge is more accurate when one can influence and interact with potential causes rather than merely observe causes and their effects (Lagnado and Sloman, 2004; Steyvers et al., 2003). Interventions are important for causal learning in the sense that they enable the differentiation of compatible causal structures through hypothesis testing (Hagmayer et al., 2007).

The interactive simulation module of IMAGE (called Multichronia) runs a computational model of a complex situation and allows the analyst to interact with this model by creating “what-if” simulations and manipulating key parameters (Lizotte et al., 2012; Rioux, Bernier, and Laurendeau, 2008). When interacting with the computational model using Multichronia, three types of actions are possible: Creating a simulation instance with new initial conditions; changing the value of a parameter at one point in time; and creating diverging simulation branches at different points in time (forming a multichronic tree, see Figure 1) to observe the impacts of different parameters on various measures of performance (MoP). For the purpose of the experiment described below, the parameters of each simulation had to be specified by the analyst. Consequently, it was not possible to simulate the model in “batch-run” mode, thus ensuring that analysts would interact with the simulation model and actively engage in hypothesis testing.

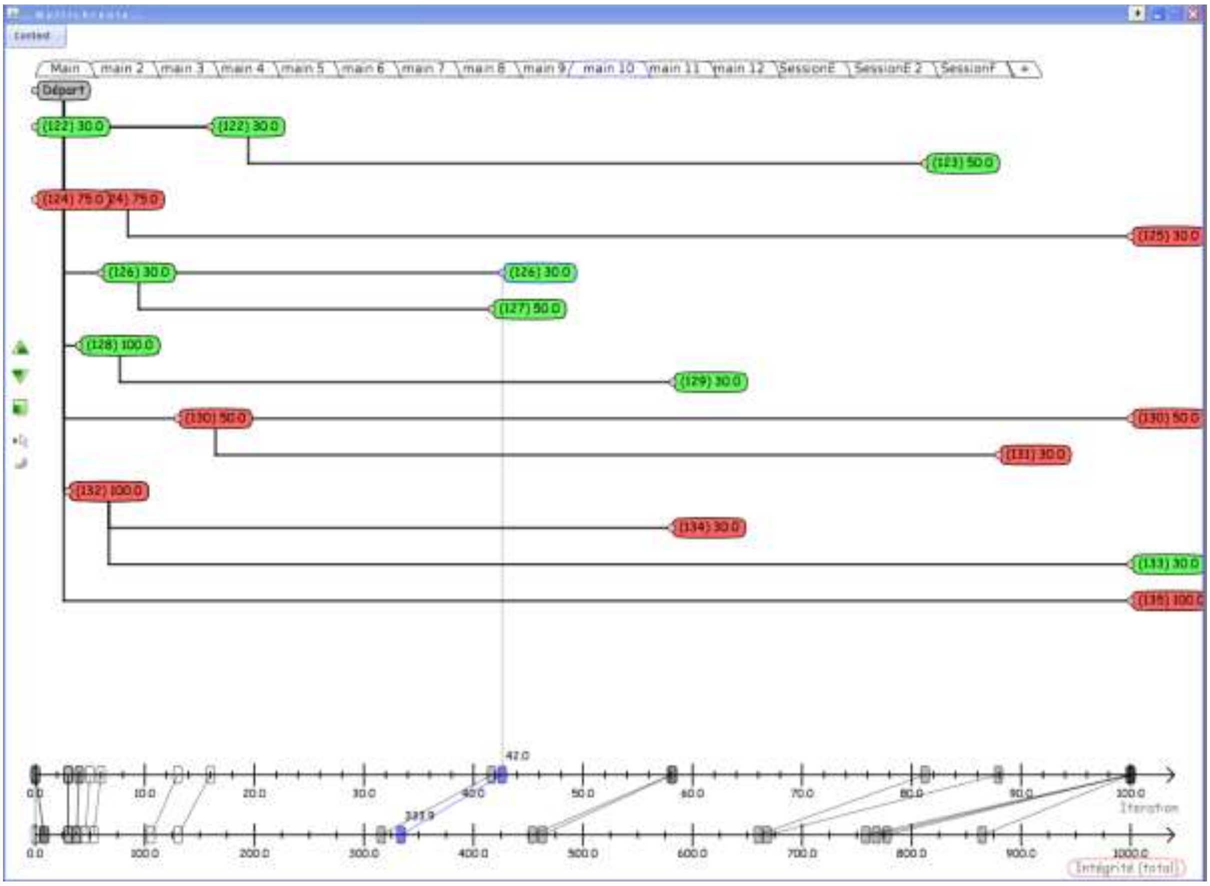

Figure 1. The interactive simulation was supported by Multichronia, a generic visual interactive simulation exploration framework providing users with a visual history of a series of simulation. Multichronia supports hypothesis-testing by allowing the creation of multichronic trees which let users simulate an instance and observe what would happen if a particular parameter was changed at any given moment in time. Whenever a parameter is changed, a new "branch" is created which visually evolves alongside its originating branch, thus making comparison between the original and the newly created instance easy for users. Branches correspond to what-if scenarios that may help understand the effects of specific changes through time.

### *1.2.2. Enhanced visualizations*

Visualization is used for various functionalities such as data aggregation (e.g., Kandel et al., 2012), coordinating multiple views, and linking different sets of data to assess relationships between dimensions (e.g., Gonzalez-Torres et al., 2013). Visually representing pre-processed data allows analysts to infer relationships or detect patterns without being constrained by cognitive limits such as the bottleneck of short-term memory (Thomas and Cook, 2005; 2006). This has been shown to improve performance in a wide variety of contexts including bioinformatics (Baehrecke et al. 2004), medicine (Tominski, Schulze-Wollgast, and Schumann, 2008), databases (Shneiderman, 2008), and e-Learning (Aguilar, Theron, and Peñalvo, 2009). Furthermore, visualizations can be improved by employing immersive virtual environments (van Dam et al., 2000) such as a Cave automatic virtual environment (CAVE; Cruz-Neira et al., 1992; Demiralp et al., 2003). A CAVE typically comprises three to six projectors arranged to display data on the walls of a room-sized cube, creating an environment that surrounds the user and provides a sense of immersion. Immersive tools can improve the identification of data clusters (Arms, Cook, and Cruz-Neira, 1999), as well as simple and complex searches (Laha et al., 2012). Such benefits of immersive virtual environments are partly explained by increased “presence” – an increased task focus resulting from the feeling of “being there” (Nash et al., 2000) – which is assumed to lead to a more sustained allocation of attentional resources and in turn, improved performance. However, the question remains as to whether a greater “presence” and more focused attention is sufficient to improve the understanding of complex situations.

IMAGE offers a toolbox of enhanced visualizations (Girardin, 2012; Lizotte et al., 2012; Mokhtari, Boivin, and Drolet, 2013; Tye-Gingras, 2011), developed using Eye-Sys software (IDV inc.) that provides the analyst with relevant feedback about his/her interventions on the model (i.e., the data generated by the simulations he/she initiated) and about the relationships between the various dimensions involved in the simulations (Figure 2). It is designed to foster analysts' understanding of the situation and support decision-making by providing feedback on the relevant dimensions, facilitating the examination of relations between those dimensions and how they interact through time. This includes key functionalities to support pattern recognition and inference, such as data aggregation, clustering, sorting and comparison (through the coordination of multiple views).

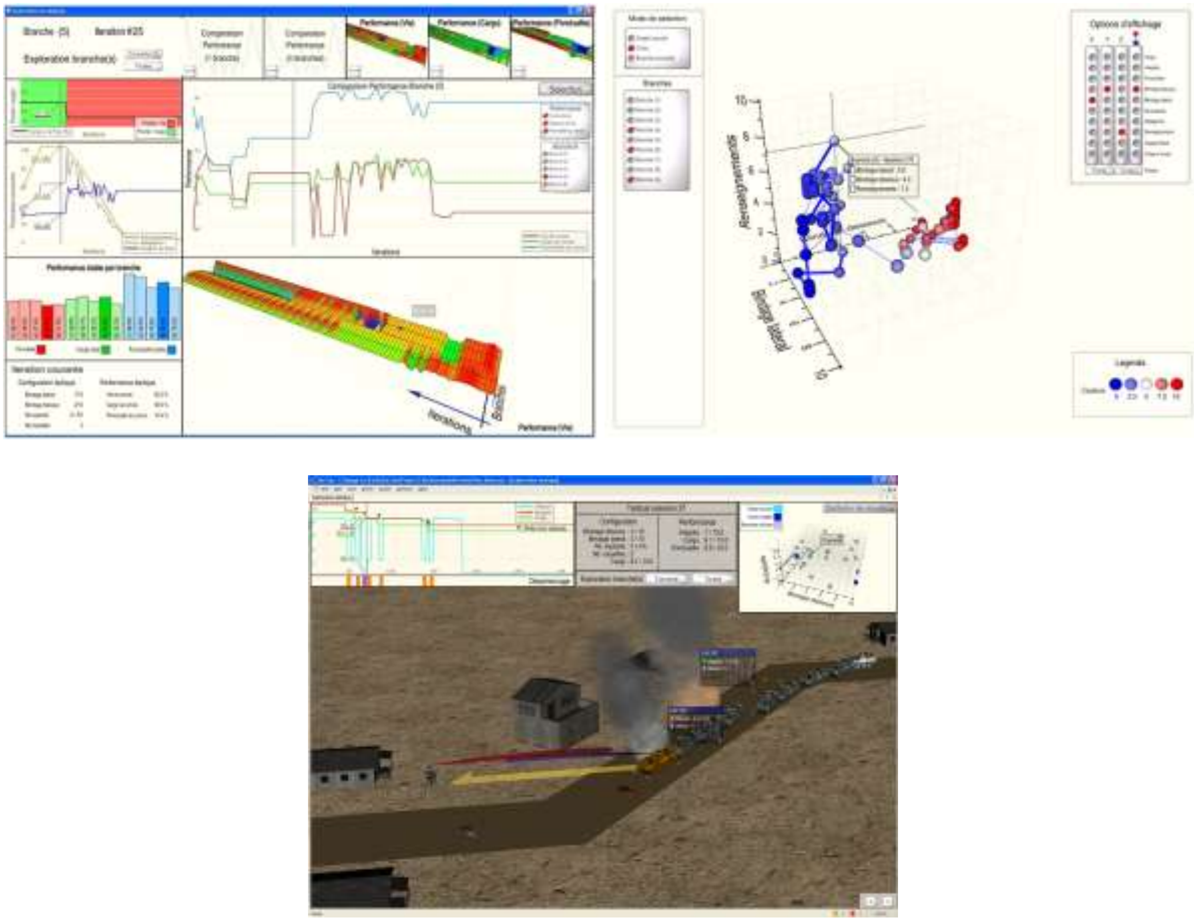

Figure 2. Analytical interactive visualizations have been developed with the Eye-Sys software package (http://www.eye-sys.com). Eye-Sys allows the gathering and manipulation of data in real-time from different sources for driving interactive, real-time visualizations. Within IMAGE, Eye-Sys provided a scientific dashboard with a variety of views aimed at facilitating the comparison of outcomes as a function of the parameter allocations. A ribbon view allowed the participant to inspect the development of MoPs across 100 iterations (top left). A cubic view allowed one to assess the impact of any three variables for each simulation (top right). The main difference between Eye-Sys in IMAGE-Desktop and IMAGE-CAVE was the presence of immersive views for the latter, whereas the former only offered traditional interactive views. Finally, a tactical view allowed analysts to see actual convoy movements (bottom).

### *1.2.3. Knowledge representation*

Internalized structural knowledge about a situation can be referred to as a mental model (Rouse and Morris 1986); a hypothetic and abstract representation of the causal structure of an environment which may vary in its degree of accuracy or proximity to reality (Gentner, 2001). Mental models of complex dynamic situations tend to be logically incomplete and oversimplified (Funke, 2001; Sterman, 1994); that is, rather than comprehensive causal networks, they are functional approximations allowing individuals to interact with a phenomenon (Jones et al., 2011). It is possible to externalize individual mental models; moreover studies show this may support their construction and maintenance by promoting a more focused search, organizing information into coherent structures, and associating with prior knowledge (Khalifa and Shen, 2005; Liu, Chen, and Chang, 2010; Stull and Mayer, 2007). Positive effects of knowledge representation have been observed in several domains of application, notably in reading comprehension (Liu et al., 2010; Ruddel and Boyle, 1989), classroom education (Jegede, Alaiyemola, and Okebukola, 1990; Trowbridge and Wandersee, 1994) and complex business problem-solving tasks (Slof et al., 2010).

IMAGE uses an external representation tool built by Defence R&D Canada (DRDC) from an existing tool, CoGUI, that was developed in the context of research efforts by Chein and Mugnier (2009) and Genest (2012). This tool called CoGUI-IMAGE (Figure 3; see also Chapter 4 of Lizotte et al., 2012), allows the user to develop conceptual graphs (Chein and Mugnier, 1992) to help express one’s understanding and iteratively develop more elaborate representations of the complex system of interest. Indeed a key goal of IMAGE is to support the development of a progressively more elaborate “comprehension model” through cycles of what-if simulations, data visualization, and knowledge synthesis.

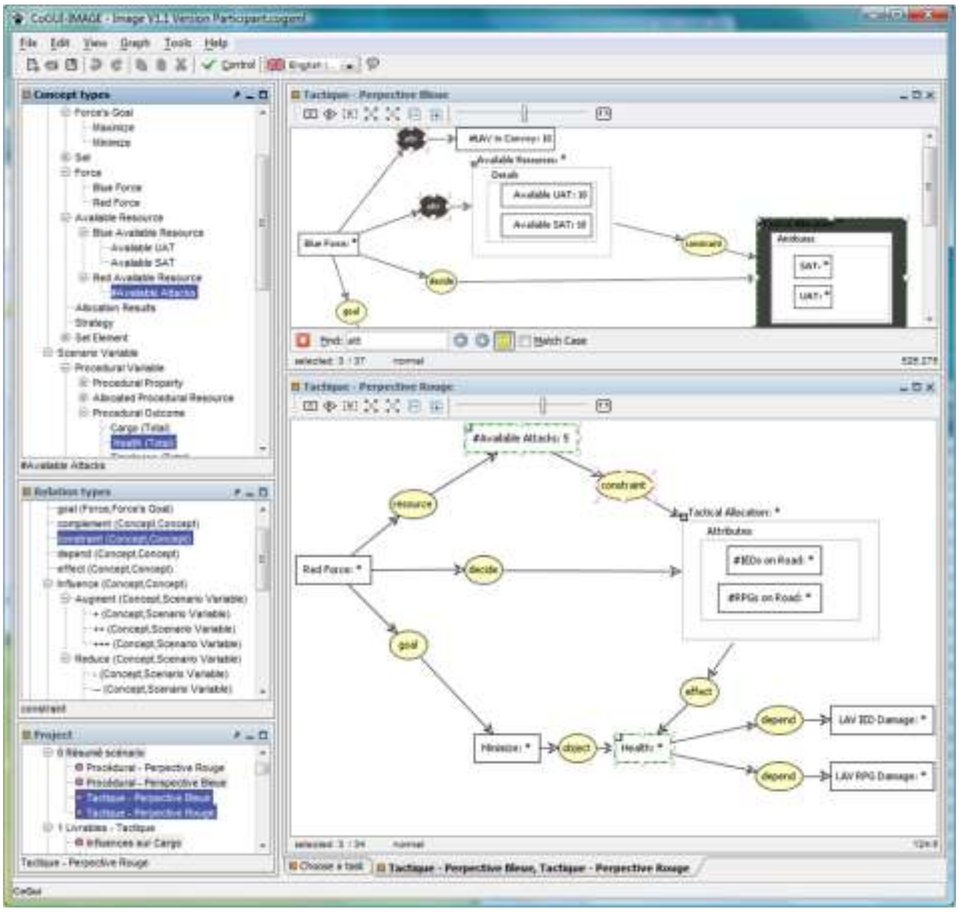


Figure 3. CoGUI-IMAGE is a tool for editing conceptual graphs. It was built by Defence R&D Canada (DRDC) from an existing tool called CoGUI, developed by the Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier.

### *1.2.4. Immersive setup*

Two versions of the IMAGE tool were implemented: IMAGE-Desktop was designed to run on a desktop computer equipped with three screens arranged in an arc (see Figure 4); and IMAGE-CAVE was designed to run in the Virtual Immersive Facility at DRDC– Valcartier (Québec, Canada) (see Figure 5). In the IMAGE-Desktop condition, the left screen displayed the knowledge representation tool, the center screen displayed the simulation tool, and the right screen displayed the enhanced visualizations of the simulation data. The functions implemented in the IMAGE-CAVE condition were equivalent to those implemented in the IMAGE-Desktop condition (i.e., interactive simulation, enhanced visualizations and knowledge representation); however, they were adapted to fit the CAVE with the goal of creating a more immersive experience. The IMAGE-CAVE had two modes of operation, changed at will by the participant. In the non-stereoscopic mode (Figure 5, first row), the left screen (not shown in Figure 5) displayed the task instructions, the center screen displayed enhanced visualizations, the right screen displayed the knowledge representation tool, and a tablet-pc fixed to the participant's chair displayed the simulation tool. In the stereoscopic mode (Figure 5, second row), four screens were dedicated to the display of enhanced visualizations, and the tablet-pc displayed the simulation tool. In this mode, participants had to wear stereoscopic glasses and controlled the visualizations with a controller wand.

Figure 4. Experimental setup in the IMAGE-Desktop condition. The screens were disposed in an arc in front of the participant. Each screen was assigned to a specific function, namely knowledge representation (left), interactive simulation (center) and enhanced visualization (right). Although not described in the present article, participants were required to wear a head-tracking device which was used for behavioural measurement and design recommendations.

Figure 5. Experimental setup for the IMAGE-CAVE condition. The CAVE operated in two distinct modes: immersive and non-immersive display. In the immersive mode, participants stood and moved around the workspace at will. An extra input device, the WAND, was added to the IMAGE toolkit in order to reproduce mouse functionalities while standing. The WAND consisted of a thumb-stick for navigating and five programmable buttons. Participants in the IMAGE-CAVE condition may also have used electronic shutter glasses (stereo-glasses) to explore two immersive interactive data visualizations called the cube view (lower left panel) and the ribbon view (lower right panel).

The assumptions behind the development of the IMAGE prosthesis are based on visual analytics. Namely, it is assumed that (a) simulations of a complex situation can help explore the parameter space and test hypotheses, (b) data resulting from those simulations can be better understood using sophisticated graphics, and (c) insights can be captured using an external knowledge representation and iteratively improved, or reframed. These three aspects are all incorporated into "IMAGE-Desktop". The IMAGE-CAVE version incorporates one additional assumption (d) implementing these functions in a CAVE may further increase benefits by augmenting the proportion of attentional resources dedicated to the execution of the task.

### *1.3. Objectives*

The study aims to validate the key assumptions of the IMAGE toolkit; first, whether the visual analytics functionalities provided by the IMAGE prosthesis help analysts understand complex situations; and second, whether immersion modulates the impact of the visual analytics tools on understanding.

## 2. Method

### *2.1. Participants*

Thirty-nine volunteers (30 men, 9 women) who had completed at least two undergraduate semesters in computer science, mathematics or operational research participated in the study. Participants received monetary compensation for their participation.

### *2.2. Task*

Participants played the role of an operations research analyst. Their main objective was to develop an understanding of a complex situation by interacting with a computational model of that situation. They were asked to gain a good enough understanding of the situation to explain its behaviors and particularities to another analyst. Specifically, participants performed two distinct tasks: an analysis and decision-making task and a structural knowledge test.

### 2.2.1. *Analysis and decision-making task*

The main task involved actively exploring the effects of potential policies and operational decisions using a computational model of a complex situation. The complex situation was first defined by a team of cognition and domain experts, then instantiated into a computational model and multi-agent simulation. The situation portrayed was the management of military convoy operations in a hostile environment. The scenario was specifically developed to encompass hallmarks of complex systems including a high number of elements, non-linear relationships, amplifying or damping feedback loops, unknown relationships, delays within the relationships between different parts of the system, counter-intuitive effects, and emergent co-evolutionary behaviour through mutual adaptation (see Bernier and Rioux, 2010). Using the Multichronia simulation tool, participants interacted with the computational model by sampling the parameter space, and manipulating the value of various parameters (e.g., armor thickness) to determine how this affected convoy mission measures of performance (MoP). Participants' goal was to find the best set of parameter values to maximize three MoPs: convoy integrity (the final health factor of the convoy after incurring damages during hostile encounters), convoy timeliness (time taken for the convoy to reach its destination) and convoy cargo (quantity of cargo brought to destination). This was not trivial as the parameter space included about six billion possibilities, thus making it impractical to exhaustively explore the entire solution space.

### 2.2.2. *Structural knowledge test*

The structural knowledge test required participants to approximate the relationships between all key variables of the model in a simplified linear fashion (i.e., by estimating the strength and direction – negative or positive – of the relationships). This was implemented through a web-based questionnaire, only available at the end of the experiment, when the digital tools and notes were no longer available. The questionnaire presented a matrix of cells allowing the participant to estimate the correlation (-1 to +1) between the variables of each row and column (Figure 6). For each pair of variables, participants had to specify the strength and direction of the linear relationship.

| | Allegiance | Intelligence | Percentage of cargo Offered | Percentage of cargo for Operations | Blue Support | Convoy Cargo | Convoy Timeliness | Convoy Integrity |
|---|---|---|---|---|---|---|---|---|
| Allegiance | | 0.75 | 0 | 0 | 0.5 | 0.375 | 0.25 | 0.375 |
| Intelligence | | | 0.25 | 0.375 | 0.125 | 0.125 | 0.25 | 0.375 |
| Percentage of Cargo Offered | | | | -0.75 | -0.625 | -0.375 | -0.5 | -0.375 |
| Percentage of Cargo for Operations | | | | | -0.625 | -0.75 | -0.625 | -0.25 |
| Blue Support | | | | | | -0.375 | -0.25 | -0.5 |
| Convoy Cargo | | | | | | | -0.5 | -0.625 |
| Convoy Timeliness | | | | | | | | -0.875 |
| Convoy Integrity | | | | | | | | |

Figure 6. The structural knowledge test consisted of a matrix with variable names in rows and columns, and cells allowing the participant to estimate the linear correlation (between -1 and +1) between the variables of each row and column.

### *2.3.Experimental design*

To assess the effect of the IMAGE toolset on structural knowledge and decision-making performance, the experiment compared a group of participants using IMAGE (i.e., the IMAGE condition which combines IMAGE-CAVE and IMAGE-Desktop users) with a group of participants using a more traditional set of tools (i.e., the baseline condition, which replicated the environment in which this type of work is usually performed). The relevance of the tools provided in the baseline condition was validated by subject matter experts (SMEs) in operations research. In the baseline condition, the left screen displayed a standard text editor (WordPad), the center screen displayed a basic simulation tool for setting model parameters and launching simulations with a limited visualization capacity (i.e., a reduced version of Multichronia essentially allowing successive what-if simulations but without the advanced multichronic tree comparison capability designed to support hypothesis testing), and the right screen displayed an Excel spreadsheet showing simulation outputs. In order to assess the potential impact of immersion when using a tool such as IMAGE, we divided the IMAGE condition into two sub-groups, one with immersion capabilities (IMAGE-CAVE) and the other without (IMAGE-Desktop). As both groups were given the IMAGE functionalities, the only difference between the two conditions was the level of immersion. Table 1 shows a breakdown of the functions provided to the participants by each version of the tool. The greyed part corresponds to the core of the IMAGE concept, which can be implemented in an immersive environment or not.

Table 1. Functions by experimental condition

| | **Functions** | **IMAGE-CAVE** | **IMAGE-Desktop** | **Baseline** |
|---|---|---|---|---|
| IMAGE Concept | (1) Interactive simulation | YES | YES | NO |
| IMAGE Concept | (2) Enhanced visualizations | YES | YES | NO |
| IMAGE Concept | (3) Knowledge representation | YES | YES | NO |
| | (4) Immersive display | YES | NO | NO |

### *2.4.Procedure*

Participants were quasi-randomly[1] assigned to one of the conditions: baseline, IMAGE-Desktop or IMAGE-CAVE. The experiment consisted of a 45-minute tutorial session, followed by two decision-making task cycles (lasting 150 minutes each, divided into two sessions of 75 minutes), ending with a web-based structural knowledge test.

[1] The unequal size of the three groups is due to a technical request to have a greater amount of open-ended feedback and system testing for the IMAGE-Desktop as well as the IMAGE-CAVE condition.

The tutorial session aimed to familiarize participants with the tools, and was completed with the help of an experimenter who ensured that each participant was proficient in the use of the tools available. The experimenter systematically prompted participants to perform all operations relevant to the task (e.g. running a simulation) making sure that each one was understood. The tutorial scenario (identical across all experimental conditions) involved the simulation of hemlock migration and its effect on the propagation of a disease; this was intentionally far-removed from the experimental scenario (convoy operations) to ensure that there was no carryover effect.

## ***2.5.Metrics***

Two measures were taken: decision-making performance and quality of structural knowledge.

### *2.5.1. Decision-making performance*

Participants were asked to separately maximize three MoPs: Integrity, Timeliness, and Cargo. For each MoP, participants had to provide the value of the parameters which would maximize its outcome. The score for each MoP was the proximity between the best possible outcome and that of the participant. Because each MoP was associated with a distinct and non-uniform distribution, some values were more frequent than others, and therefore potentially easier to find by chance. In order to assign a score which rewarded the participant proportionally to the difficulty of attaining a given score, raw scores obtained by participants were normalized using a measure taking into account the three distributions. To obtain these distributions, 750,000 simulations were launched according to a pseudo-random exploration of the parameter space using the scrambled Halton sequence approach (Bhat 2003). From these distributions, it was found that high scores for convoy timeliness were the least difficult to obtain, since higher values could be reached by a multiplicity of parameter combinations.

High convoy integrity scores were more difficult to achieve, and high values for convoy cargo were the most difficult. For this reason, the scores were normalized based on their distribution, ranging from 0 (lowest performance) to 1 (highest performance). Decision-making performance was the average of the normalized scores for the three MoPs from both decision-making task cycles.

In the present study, good decision-making performance could be achieved through an effective sampling of the parameter space (i.e., an efficient search for optimal parameters). The exploration of the parameter space and the acquisition of structural knowledge (hypothesized to lead to better sampling) constitute the essence of the present experimental task and essentially refer to a cyclic process of hypothesis testing and analytic learning. It is hypothesized that participants who develop a better understanding of the relations between the parameters and the system's variables will be more inclined to sample the parameter space adequately and consequently to achieve a higher decision-making performance.

### *2.5.2. Quality of structural knowledge*

The web-based questionnaire explicitly probed participants' knowledge of the relationships between key variables in the computational model of the complex situation. The quality of participants' structural knowledge was estimated by taking into account both how close to ideal and how far from random their responses were. The proximity to the ideal and to a random solution was calculated for each participant by calculating a normalized distance between participants' responses and the two reference matrices. The ideal reference matrix was obtained by deriving the correlation values between variables using a representative sample of the parameter space. The random reference matrix was equivalent to a matrix filled with zeros (i.e. no strength and no direction). The quality of structural knowledge was equal to the average error with the ideal matrix divided by the average error with the random reference matrix. The value of this metric favours participants that are both close to the ideal solution and further from random. While this metric can vary from zero to infinity, a value of one corresponds to a response matrix that is equidistant from both the ideal and the random matrices. A value above one means that responses are closer to the ideal matrix, while a value below one means that responses are closer to the random matrix. A complementary measure of structural knowledge could in principle be derived from the knowledge representations of participants (developed in either WordPad or CoGUI), yet comparing different types of representations either quantitatively or qualitatively comes with methodological challenges that are beyond the scope of the present work. Here, the focus lies on assessing the impact of IMAGE as a whole, where knowledge representation forms an integral part of the hypothesis testing and iterative learning process.

## 3. Results

### *3.1. Intergroup analyses*

In a first step, differences between the experimental groups were assessed in terms of decision-making performance and structural knowledge. A statistically significant difference on decision-making performance was observed between the IMAGE and baseline conditions $t(37)= 2.07$, $p = .045$. Participants in the IMAGE conditions (M = .80, SD = .07) obtained higher mean decision-making performance than participants in the baseline condition (M = .74, SD = .09) (Figure 7). This represents an average increase of 6% in terms of decision-making performance when using IMAGE tools in comparison with baseline tools.

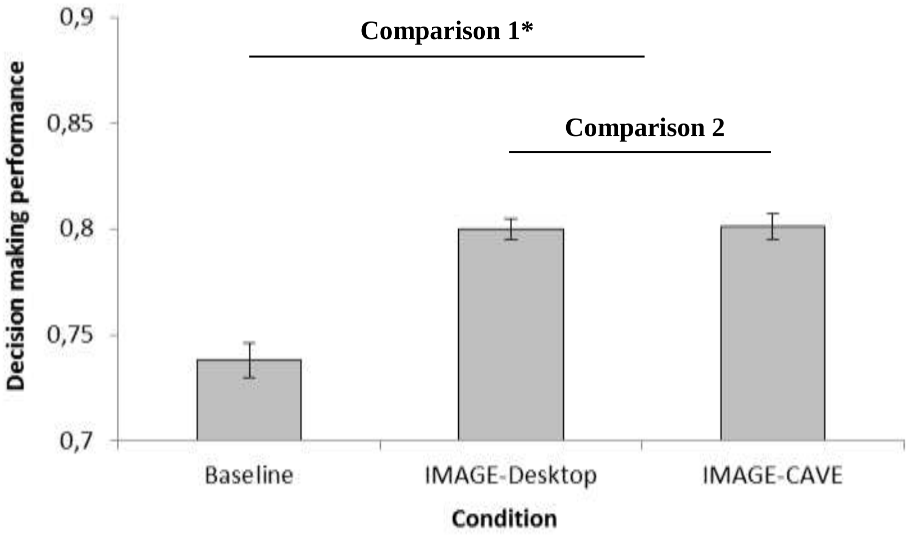


Figure 7. Decision-making performance by experimental condition. Error bars represent mean squared error.

No statistically significant difference on decision-making performance was observed between the IMAGE-desktop and IMAGE-CAVE conditions $t(28) < 1$. Participants in the IMAGE-desktop condition (M = .80, SD = .07) did not obtain a higher mean score than participants in the IMAGE-CAVE condition (M = .80, SD = .08).

Mean quality of structural knowledge as measured by the covariation matrix was not significantly different across experimental conditions $t(37) < 1$. Participants in the IMAGE condition (M = .99, SD = .06) did not exhibit better structural knowledge than participants in the baseline condition (M = .97, SD = .08). Furthermore, structural knowledge was not significantly different between the two IMAGE conditions $t(28) < 1$. Indeed, participants in the IMAGE-CAVE group (M = 1.00, SD = .09) did not differ from participants in the IMAGE-Desktop condition (M = .99, SD = .04) in terms of quality of structural knowledge (see Figure 8). For the various non-significant tests reported above, a power analysis shows that it would potentially require between 115 and 6000 participants for the observed differences to be statistically significant. For all practical purposes, we can safely conclude there was no reliable *operationally relevant* difference for these last three comparisons.

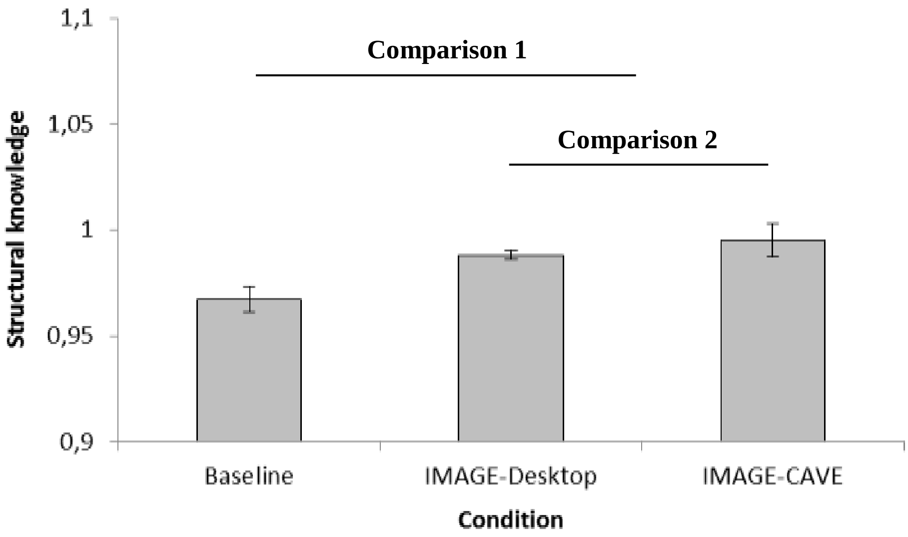


Figure 8. Structural knowledge by experimental condition. Error bars represent mean squared error.

### *3.2. Regression analyses*

In a second step, we performed three linear regressions to assess the relation between the number of what-if simulations and (a) decision making performance and (b) structural knowledge, as well as between structural knowledge and decision making performance. Since there was no significant difference across experimental conditions in terms of number of what-if simulations $F(36) = 1.451$, *N.S*, it was decided to pool results from experimental conditions together to increase statistical power. Standard linear regressions were then carried out between, on the one hand (a) decision making performance (M = .82, SD = .09) as the dependent variable (DV) and number of what-if simulations (M = 79.77, SD = 33.87) as the independent variable (IV) and on the other hand between (b) quality of structural knowledge (M = 98.55, SD = .07) as the DV and number of what-if simulations as the IV. Each of the two linear regressions is described in its own section below. Analyses were performed using SPSS REGRESSION and SPSS EXPLORE for the evaluation of assumptions. All assumptions, including normality, linearity, homoscedasticity of residuals, and the absence of outliers, were met.

### *3.2.1. Number of what-if simulations and decision making performance*

Table 2 displays the correlation between variables, the unstandardized regression coefficient (*B*) and intercept, the standardized regression coefficient (β), $R^2$ and adjusted $R^2$. *R* for regression was significantly different from zero, $F(1, 37) = 4.396$, $p = .043$, with $R^2$ at .106. The adjusted $R^2$ value of .082 indicates that approximately 8% of the variability of decision making performance can be predicted by the number of what-if simulations.

Table 2. Linear regression of number of simulations on decision making performance

| **Variables** | **Decision making performance (DV)** | **Number of what-if simulations (IV)** | ***B*** | **β** |
|---|---|---|---|---|
| Number of whaf-if simulations | .326 | Intercept = 0.72 | 0.001* | 0.326 |
| Means | 0.786 | 79.77 | | |
| Standard deviations | 0.081 | 33.87 | $R^2$ = .106<br>Adjusted $R^2$ = .082<br>$R$ = .326* | |

* $p = .043$

The direction of the relationship suggests that a greater amount of interaction with the model (i.e., launching/modifying simulation runs) is associated with greater decision making performance.

### *3.2.2. Number of what-if simulations and quality of structural knowledge*

Table 3 displays the correlation between variables, the unstandardized regression coefficient (*B*) and intercept, the standardized regression coefficient (β), $R^2$ and adjusted $R^2$. *R* for regression was significantly different from zero, $F(1, 37) = 5.319$, $p = .027$, with $R^2$ at .126.

The adjusted $R^2$ value of .102 indicates that approximately 10% of the variability of quality of structural knowledge can be predicted by the number of what-if simulations.

Table 3. Linear regression of number of simulations on quality of structural knowledge

| **Variables** | **Structural knowledge (DV)** | **Number of what-if simulations (IV)** | ***B*** | **β** |
|---|---|---|---|---|
| Number of what-if simulations | .355 | | 0.001* | 0.355 |
| | | Intercept = 0.93 | | |
| Means | 0.986 | 79.77 | | |
| Standard deviations | 0.067 | 33.87 | | $R^2$= .126 |
| | | | | Adjusted $R^2$= .102 |
| | | | | $R$ = .355* |

* $p$ = .027

The direction of the relationship suggests that a greater number of what-if simulations is associated with greater structural knowledge. An additional regression was carried out between decision making performance and quality of structural knowledge. The regression was not statistically significant $F(1, 37) = 4.02$, *N.S.*

## 4. Discussion

This study investigated whether IMAGE, a cognitive prosthesis to support analytical reasoning, could increase analysts' understanding of a complex system and whether its immersive capability could further increase structural knowledge acquisition and decision-making performance. While a variety of analytical tools have been designed to support understanding and planning in complex domains (e.g., Allen, Corpac, and Frisbie, 2006; Chappell et al., 2004; Chen and Lee, 2003; Langton and Das, 2007; Surdu and Kitka, 2008; Vester, 2007), the effectiveness of such tools is not often investigated in a systematic manner using the experimental research method (e.g., Lafond et al., 2012; Lerch and Harter, 2001).

Results show that providing a cognitive prosthesis that incorporates interactive simulation, enhanced visualization and knowledge representation, can improve decision-making performance but does not appear to increase the accuracy of structural knowledge. Importantly, this finding suggests that the cognitive tool was indeed useful as a cognitive “prosthesis” in the sense that benefits emerged when operating as a *joint cognitive system* (i.e. a human-tool dyad), however it did not foster a better understanding of the situation more generally (i.e., IMAGE did not lead to better internal mental models as assessed by the matrix test). This is in line with the view that human analysts should collaborate with such cognitive tools, whereby each component, human or machine, performs complementary functions towards the achievement of a common goal (see Arias-Hernandez et al., 2012). For instance, in the context of the present study, IMAGE performs functions like data aggregation (by consolidating all data into visual representations), and simulation (by running the computational model), while the analyst’s role is to detect patterns and make decisions about the parameters that need to be acted upon.

Our results also show that the addition of immersion capabilities in the IMAGE-CAVE subgroup did not lead to an increase in decision-making performance nor to an increase in the quality of structural knowledge compared with the IMAGE-Desktop group. One possible explanation for this finding could be the nature of the complex situation which was mostly conceptual rather than spatial. Indeed, previous research on immersive 3D environments has shown that most benefits associated with such are demonstrated with inherently spatial tasks such as protein folding (Férey, Nelson, Martin et al., 2009). Conversely, research pertaining to the visualization of multidimensional statistical data has obtained mixed results (e.g., Arns, Cook, and Cruz-Neira, 1999).

We offer five potential (but not exclusive) explanations for the observed lack of impact of IMAGE on structural knowledge. First, it may have been that the augmented toolset provided was simply no more helpful than the baseline tools to foster development of a long-term representation of the system. Second, our measure of structural knowledge assumes linear relationships between the different situation variables (to estimate the overall direction and strength of relationships) and may consequently underestimate participants' knowledge if the latter is nonlinear. Still, we argue that our measure should be sensitive enough to detect differences in understanding, since linear estimation of non-linear relations has been shown to account for a significant proportion of variance in most cases (Karelaia and Hogarth, 2008). Third, the time allocated for creating a good internal representation of the structure of the system may have been too short. Indeed, the combinatorial complexity of the system is very high (with over 6 billion possible combinations of parameters), and the average number of simulated instances was very low. Although this was sufficient for relatively good performance on the decision-making task, one could argue that this is greatly insufficient to create a valid internal representation of the structure of the system. Fourth, this finding could also be explained by a ceiling effect caused by fundamental human cognitive limits (e.g., Halford et al., 2005). Indeed, humans may be unable to acquire and retain more than a fraction of the structural knowledge necessary to fully understand complex systems. Finally, instance based learning theory (Gonzalez, Lerch, and Lebiere, 2003) would suggest that people may have acquired procedural knowledge on how to act using an interactive tool such as IMAGE, but not necessarily causal knowledge (either explicit or implicit) about the structure of the system. The observed relation between the number of simulation runs and decision making performance, as well as the lack of a significant relation between structural knowledge and decision making performance both make sense from the perspective of instance based learning theory.

The current index of structural knowledge shows that acquired knowledge was equally close to the ideal solution as it was to the random allocation of values, leaving considerable room for improvement. In the context of the current study, such a result clearly illustrates the complexity of the simulated situation and the difficulty associated with its representation. Our finding that the number of simulation runs is related to structural knowledge (but not necessarily causally) may lead to the hypothesis that promoting a greater exploration of the parameter space would help users achieve better structural knowledge. More research is needed to identify the factors that influence the formation of a good representation of a complex system such as that simulated in the present study (see Gary and Wood, 2007). In supporting the development of structural knowledge, one factor that deserves much further investigation is the potential benefit of computational tools that keep the human in the loop and provide a means for hypothesis testing.

Leading models of sensemaking and situational awareness (e.g., Endsley, 1995; Thomas and Cook, 2005) and empirical work (e.g., Goode and Beckmann, 2010) suggest that comprehension is a key factor underlying effective decision-making, yet in the current study there was no general relationship between structural knowledge and decision-making performance. One might question then, whether this assumed relationship between structural knowledge and decision-making still holds true when dealing with highly complex problems. Decision-making performance was supported by the use of tools (all participants, even those in the baseline condition, used tools to support their analysis and decision-making), whereas no tool was available for the structural knowledge test. Perhaps it is not that surprising then that un-aided cognitive performance showed no relation to cognitive performance with tools. This points back to the notion of a cognitive prosthesis, and the idea that in complex domains cognitive artifacts are the “things that make us smart” (Norman, 1993), allowing us to be effective in ways that we could not possibly be without our ingenious devices.

In terms of practical implications, our pattern of results reinforces the relevance of developing joint cognitive systems in the effort to augment comprehension of complex situations, but also highlights the limitations of this approach. Joint cognitive systems are widely used in the support of performing complex tasks including air traffic control (e.g. Harris 2013; Hollnagel 2007), emergency and crisis management (e.g., Furniss and Blanford, 2006; Ntuen et al., 2006), warehouse management (Accorsi, Manzini, and Maranesi, 2013), and policy making (Ntuen, Park, and Gwang-Myung, 2010; Parsons and Sedig, 2014). Our results also suggest that joint cognitive systems may be insufficient for human users to develop a good internal representation of a complex situation. This is critical as most of these systems would usually be used by analysts who would have to share their comprehension with the actual decision-maker afterwards. To facilitate this process, our results suggest that the design of a joint cognitive system should leave the control in the hands of the analysts. Although many systems already support such a feature (e.g. Accorsi, Manzini, and Maranesi, 2013), still many simulations allow little or no control to the analysts over the sampling of the parameter space (see Brandenburg et al., 2013). Overall, by stressing further the importance of interactivity within the visual analytics framework, this study underlines the importance of integrating human factors considerations in the design of joint cognitive systems.

## 5. Acknowledgements

Research funding for this project was provided by a contract from DRDC Valcartier (via the Technological Investment Fund) and a grant from the Natural Science and Engineering Research Council of Canada to Sébastien Tremblay. Also, Patrick Jeuniaux received a postdoctoral fellowship from MITACS. We are thankful to Michel Lizotte, François Bernier, Marielle Mokhtari, Éric Boivin, Michel B. DuCharme and Denis Poussart for their guidance and support throughout the project. Pictures of the IMAGE and baseline tools are the courtesy of DRDC Valcartier Research Centre.

## 6. References

Accorsi, R., Manzini, R., Maranesi, F., 2014. A decision-support system for the design and management of warehousing systems. Computers in Industry, 65 (1), 175–186. doi:10.1016/j.compind.2013.08.007

Aguilar, D.A.G., Therón, R., Peñalvo, F.G., 2009. Semantic Spiral Timelines Used as Support for e-Learning. Journal of Universal Computer Science 15 (7), 1526–1545.

Agyepong, I.A., Kodua, A., Adjei, S., Adam, T., 2012. When ‘solutions of yesterday become problems of today’: crisis-ridden decision making in a complex adaptive system (CAS)—the additional duty hours allowance in Ghana. Health policy and planning, 27 (4), iv20–iv31. doi:10.1093/heapol/czs083

Allen, J.G., Corpac, P.S., Frisbie, K.R., 2006. Integrated battle command program: Decision support tools for planning and conducting unified action campaigns in complex contingencies. Defense Advanced Research Projects Agency, Arlington VA.

Arias-Hernandez, R., Green, T.M., Fisher, B., 2012. From cognitive amplifiers to cognitive prostheses: Understandings of the material basis of cognition in visual analytics. Interdisciplinary science reviews 37(1): 4-18. doi:10.1179/0308018812Z.0000000001

Arndt, H., 2007. Using system dynamics-based learning environments to enhance system thinking. Conference Proceedings, The 2007 International Conference of the System Dynamics Society and 50th Anniversary Celebration.

Arms, L., Cook, D., Cruz-Neira, C., 1999. The benefits of statistical visualization in an immersive environment. Conference Proceedings, Virtual Reality, IEEE. doi: 10.1109/VR.1999.756938

Ashby, W.R., 1968. Variety, constraint, and the law of requisite variety. In W. Buckley (ed.), Modern Systems Research for the Behavioral Scientist, Chicago, IL: Aldine Publishing Co.

Baehrecke, E.H., Dang, N., Babaria, K., Shneiderman, B., 2004. Visualization and analysis of microarray and gene ontology data with treemaps. BMC bioinformatics, 5 (1), 84. doi:10.1186/1471-2105-5-84

Bakken, B.T., Gilljam, M., 2003. Dynamic intuition in military command and control: why it is important, and how it should be developed. Cognition, technology & work, 5 (3), 197–205. doi: 10.1007/s10111-003-0123-1

Bernier, F., Rioux, F., 2010. Convoy scenario for complexity study: A coevolutionary perspective on convoy against insurgents. DRDC-VALCARTIER-TM-2010-150.

Betsch, T., Fiedler, K., Brinkmann, J., 1998. Behavioral routines in decision making: The effects of novelty in task presentation and time pressure on routine maintenance and deviation. European Journal of Social Psychology, 28 (6), 861–878. doi: 10.1002/(SICI)1099-0992(1998110)28:6<861::AID-EJSP899>3.0.CO;2-D

Betsch, T., Haberstroh, S., Glöckner, A., Haar, T., Fiedler, K., 2001. The effects of routine strength on adaptation and information search in recurrent decision making. Organizational behavior and human decision processes, 84 (1), 23–53. doi.10.1006/obhd.2000.2916

Betsch, T., Haberstroh, S., Molter, B., Glöckner, A., 2004. Oops, I did it again—Relapse errors in routinized decision making. Organizational behavior and human decision processes, 93 (1), 62–74. doi.10.1016/j.obhdp.2003.09.002

Bhat, C.R., 2003. Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. Transportation Research Part B: Methodological, 37 (9), 837–855. doi:10.1016/S0191-2615(02)00090-5

Blech, C., Funke, J., 2005. Dynamis review: An overview about applications of the Dynamis approach in cognitive psychology. URL (31.07. 2006): http://www. die-bonn. de/esprid/dokumente/doc-2005/blech05_01. pdf.

Brandenburg, M., Govindan, K., Sarkis, J., Seuring, S., 2014. Quantitative models for sustainable supply chain management: Developments and directions. European Journal of Operational Research, 233 (2), 299–312. doi:10.1016/j.ejor.2013.09.032

Bryant, D.J., Webb, R.D.G., McCann, C., 2003. Synthesizing two approaches to decision making in command and control. Canadian Military Journal, 4 (1), 29–34.

Busenitz, L.W., Barney, J.B., 1997. Differences between entrepreneurs and managers in large organizations: Biases and heuristics in strategic decision-making. Journal of business venturing, 12 (1), 9–30. doi:10.1016/S0883-9026(96)00003-1

Chappell, A.R., Cowell, A.J., Thurman, D.A., Thomson, J.R., 2004. Supporting mutual understanding in a visual dialogue between analyst and computer. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 48 (3), 376–380. SAGE Publications. doi: 10.1177/154193120404800322

Chein, M., Mugnier, M.-L., 2008. Graph-based knowledge representation: computational foundations of conceptual graphs. Springer, London.

Chen, J.Q., Lee, S.M., 2003. An exploratory cognitive DSS for strategic decision making. Decision support systems, 36 (2), 147–160. doi. 10.1016/S0167-9236(02)00139-2

Churchman, C.W., 1967. Wicked problems. Management Science, 14 (4), 141–142.

Cohen, E.A., 2012.Military misfortunes: The anatomy of failure in war. New York: Free Press.

Cooper, J.R., 2005. Curing Analytic Pathologies: Pathways to Improved Intelligence Analysis. Washington, DC: Center for the Study of Intelligence.

Cruz-Neira, C., Sandin, D.J., DeFanti, T.A., Kenyon, R.V., Hart, J.C., 1992. The CAVE: audio visual experience automatic virtual environment. Communications of the ACM 35(6): 64-72. doi:10.1145/129888.129892

Davis, M.A., Curtis, M.B., Tschetter, J.D., 2003. Evaluating cognitive training outcomes: Validity and utility of structural knowledge assessment. Journal of Business and Psychology, 18 (2), 191–206. doi. 10.1023/A:1027397031207

Demiralp, C., Laidlaw, D.H., Jackson, C., Keefe, D., Zhang, S., 2003. Subjective usefulness of CAVE and fish tank VR display systems for a scientific visualization application. In Proceedings of the 14th IEEE Visualization 2003 (VIS'03). IEEE Computer Society. doi:10.1109/VIS.2003.10019

Diehl, E., Sterman, J.D., 1995. Effects of feedback complexity on dynamic decision making. Organizational Behavior and Human Decision Processes, 62 (2), 198–215. doi. 10.1006/obhd.1995.1043

Dörner, D., 1996. The logic of failure: Recognizing and avoiding error in complex situations. Reading, MA: Addison-Wesley.

Drigas, A., Koukianakis, L., Papagerasimou, Y., 2011. Towards an ICT-based psychology: E-psychology."Computers in Human Behavior, 27 (4), 1416–1423. doi:10.1016/j.chb.2010.07.045

Endsley, M.R., 1995. Toward a theory of situation awareness in dynamic systems. Human Factors, 37 (1), 32–64. doi. 10.1518/001872095779049543

Férey, N., Nelson, J., Martin, C., Picinali, L., Bouyer, G., Tek, A., Bourdot, P., Burkhardt, J.M., Katz, B.F.G., Ammi, M., Etchebest, C., Autin, L., 2009. Multisensory VR interaction for protein-docking in the CoRSAIRe project. Virtual Reality, 13 (4), 273–293. doi. 10.1007/s10055-009-0136-z

Fildes, R., Goodwin, P., 2007. Against your better judgment? How organizations can improve their use of management judgment in forecasting. Interfaces, 37 (6), 570–576. doi:10.1287/inte.1070.0309

Forrester, J.W. 1993. System dynamics and the lessons of 35 years. In A systems-based approach to policymaking (pp. 199–240). Springer US.

Funke, J., Frensch, P.A., 1995. Complex problem solving: The European perspective.

Funke, J., 2001. Dynamic systems as tools for analysing human judgement. Thinking & reasoning, 7 (1), 69–89. doi:10.1080/13546780042000046

Furniss, D., Blandford, A., 2006. Understanding Emergency Medical Dispatch in terms of Distributed Cognition: a case study. Ergonomics, 49 (12-13), 1174-1203. doi:10.1080/00140130600612663

Gary, M., Wood, R., 2007. Testing the effects of a system dynamics decision aid on mental model accuracy and performance on dynamic decision making tasks. Proceedings of the 25th international conference of the system dynamics society, Boston, USA.

Genest, D., 2012. Cogitant: Conceptual Graphs Integrated Tools allowing Nested Typed graphs 5. Retrieved June 20, 2012, from http://cogitant.sourceforge.net/cogitant_html/index.html

Gentner, D., 2002. Psychology of Mental Models. In N. J. Smelser & P. B. Bates (Eds.), International Encyclopedia of the Social and Behavioral Sciences (pp. 9683–9687). Amsterdam: Elsevier Science.

Gigerenzer, G., Gaissmaier, W., 2011. Heuristic decision making.Annual review of psychology 62, 451–482. doi:10.1146/annurev-psych-120709-145346

Gigerenzer, G., Selten, R., 2001. Rethinking rationality. Bounded rationality: The adaptive toolbox: 1–12.

Girardin, M., 2012. Visualisation d'information pour l'aide à la compréhension de situations complexes. Mémoire de maîtrise, Université Laval, Québec, Canada.

Gonzalez, C., Lerch, J.F.,  Lebiere, C. (2003). Instance-based learning in dynamic decision making. Cognitive Science, 27 (4), 591–635.

Gonzalez, C., Vanyukov, P., Martin, M.K., 2005. The use of microworlds to study dynamic decision making. Computers in Human Behavior, 21 (2), 273–286.doi. 10.1016/j.chb.2004.02.014

González-Torres, A., Peñalvo, F.G., Therón, R., 2013. Human–computer interaction in evolutionary visual software analytics. Computers in Human Behavior,  29 (2), 486–495. doi. 10.1016/j.chb.2012.01.013

Goode, N., Beckmann, J.F., 2010. You need to know: There is a causal relationship between structural knowledge and control performance in complex problem solving tasks. Intelligence, 38 (3), 345–352. doi:10.1016/j.intell.2010.01.001

Goodwin, P., 2005. How to integrate management judgment with statistical forecasts. Foresight: The International Journal of Applied Forecasting, 1, 8–12.

Granlund, R., 2003. Monitoring experiences from command and control research with the C3Fire microworld. Cognition, Technology & Work, 5 (3), 183–190. doi:10.1007/s10111-003-0129-8

Hagmayer, Y., Sloman, S.A., Lagnado, D.A., Waldmann, M.R., 2007. Causal reasoning through intervention. In A. Gopnik & L. Schulz (Eds.), Causal learning: Psychology, philosophy, and computation (pp. 86–100). Oxford: Oxford University Press.

Halford, G.S., Baker, R., McCredden, J.E., Bain, J.D., 2005. How many variables can humans process?. Psychological science, 16 (1), 70–76. doi:10.1111/j.0956-7976.2005.00782.x

Harris, D., 2013. Distributed cognition in flight operations. In D. Harris (Ed.), Engineering Psychology and Cognitive Ergonomics. Applications and Services (pp. 125–133). Springer.

Heuer, R.J., 1999. The Psychology of Intelligence Analysis. Washington, DC: Center for the Study of Intelligence.

Hollnagel, E., 2007. Flight decks and free flight: Where are the system boundaries? Applied ergonomics, 38 (4), 409–416.

Jegede, O.J., Alaiyemola, F.F., Okebukola, P.A.O., 1990. The effect of concept mapping on students' anxiety and achievement in biology. Journal of research in science teaching, 27 (10 ), 951–960. doi. 10.1002/tea.3660271004

Johnston, R., 2005. Analytic Culture in the US Intelligence Community: an Ethnographic Study. Washington, DC: Center for the Study of Intelligence.

Jones, N., Ross, H., Lyman, T., Perez, P., Leitch, A., 2011. Mental Models: An Interdisciplinary Synthesis of Theory and Methods. Ecology and Society, 16 (1), 46.

Kandel, S., Parikh, R., Paepcke, A., Hellerstein, J.M., Heer, J., 2012. Profiler: Integrated statistical analysis and visualization for data quality assessment. In Proceedings of the International Working Conference on Advanced Visual Interfaces, pp. 547–554. ACM.

Karakul, M., Qudrat-Ullah, H., 2008. How to improve dynamic decision making? Practice and promise. Complex Decision Making. Springer Berlin Heidelberg, 3–24.

Karelaia, N., Hogarth, R.M., 2008. Determinants of linear judgment: a meta-analysis of lens model studies. Psychological Bulletin, 134 (3), 404. doi:10.1037/0033-2909.134.3.404

Kerren, A., Schreiber, F., 2012. Toward the role of interaction in visual analytics. Proceedings of the Winter Simulation Conference. Winter Simulation Conference.

Khalifa, M., Ning Shen, K., 2005. Effects of knowledge representation on knowledge acquisition and problem solving. Proceedings of the 2nd International Conference on Intellectual Capital, Knowledge Management and Organisational Learning (ICICKM 2005). Academic Conferences Limited.

Keim, D.A., Mansmann, F., Schneidewind, J., Thomas, J., Ziegler, H., 2008. Visual analytics: Scope and challenges. Springer Berlin Heidelberg.

Klein, G., Moon, B.M., Hoffman, R.R., 2006. Making Sense of Sensemaking 1: Alternative Perspectives. IEEE intelligent systems, 21(4), 70-73. doi:10.1109/MIS.2006.75

Lafond, D., DuCharme, M.B., Gagnon, J.-F., Tremblay, S., 2012. Support requirements for cognitive readiness in complex operations. Journal of Cognitive Engineering and Decision Making, 6 (4), 393-426. doi. 10.1177/1555343412446193

Lagnado, D.A., Sloman, S., 2004. The advantage of timely intervention. Journal of Experimental Psychology: Learning, Memory, and Cognition, 30 (4), 856. doi. 10.1037/0278-7393.30.4.856

Laha, B., Sensharma, K., Schiffbauer, J.D., Bowman, D.A., 2012. Effects of immersion on visual analysis of volume data. IEEE Transactions on Visualization and Computer Graphics, 18 (4), 597-606. doi.10.1109/TVCG.2012.42

Langton, J., Das, S., 2007. A framework for building and reasoning with adaptive and interoperable PMESII models. (Air Force Research Laboratory Tech. Rep. AFRL-RI-RS-TR-2007-241). Cambridge, MA: Charles River Analytics.

Lerch, F.J., Harter, D.E., 2001. Cognitive support for real-time dynamic decision making. Information systems research 12 (1), 63–82. doi: 10.1287/isre.12.1.63.9717

Liu, P.L., Chen, C.-J., Chang, Y.J., 2010. Effects of a computer-assisted concept mapping learning strategy on EFL college students' English reading comprehension. Computers & Education, 54 (2), 436–445. doi:10.1016/j.compedu.2009.08.027

Lizotte, M., Bernier, F., Mokhtari, M., Boivin, E., 2012. IMAGE Final Report - An Interactive Computer-Aided Cognition Capability (Technical report No. TR 2012-397) (p. 132). Québec, Canada: DRDC Valcartier.

Milrad, M., 2004. Learning with Models and Learning by Modelling: Exploring the Role of Multiple Representations Using Computational Media. Joensuu, Finland.

Mitchell, D., Massoud, T.G., 2009. Anatomy of Failure: Bush's Decision-Making Process and the Iraq War. Foreign Policy Analysis, 5 (3), 265-286. doi. 10.1111/j.1743-8594.2009.00093.x

Mokhtari, M., Boivin, E., Drolet, F., 2013. Visualization to Augment Comprehension Models of a Complex Phenomenon : IMAGE V1 Exploration Module (Technical memorandum No. TM 2013-482) (p. 86). Québec, Canada: DRDC Valcartier.

Nash, E.B., Edwards, G.W., Thompson, J.A., Barfield, W., 2000. A review of presence and performance in virtual environments.International Journal of human-computer Interaction , 12 (1), 1–41. doi. 10.1207/S15327590IJHC1201_1

Naval Postgraduate School. 2000. Operations research curricula produce decision-makers. Naval Postgraduate School Research, 10 (2), 6–46.

Norman, D.A., 1993. Things that make us smart: Defending human attributes in the age of the machine. Reading. MA: Addison-Wesley Publishing.

Ntuen, C.A., Balogun, O., Boyle, E., Turner, A., 2006. Supporting command and control training functions in the emergency management domain using cognitive systems engineering. Ergonomics, 49 (12-13). 1415-1436. doi:10.1080/00140130600613083

Ntuen, C.A., Park, E.H., Gwang-Myung, K., 2010. Designing an information visualization tool for sensemaking. International Journal of Human–Computer Interaction, 26 (2-3), 189–205. doi:10.1080/10447310903498825

Özesmi, U., Özesmi, S.L., 2004. Ecological models based on people's knowledge: a multi-step fuzzy cognitive mapping approach. Ecological Modelling, 176 (1), 43–64. doi:10.1016/j.ecolmodel.2003.10.027

Osman, M., 2010a. Controlling uncertainty: A review of human behavior in complex dynamic environments. Psychological Bulletin, 136 (65-86). doi. 10.1037/a0017815

Osman, M., 2010b. Controlling uncertainty: How we learn about, and make decisions in complex worlds. Oxford: Wiley-Blackwells Publishing.

Parsons, P., Sedig, K., 2014. Handbook of Human Centric Visualization. In W. Huang (Ed.), (pp. 693–715). New York, NY: Springer New York. doi:10.1007/978-1-4614-7485-2

Quesada, J., Kintsch, W., Gomez, E., 2005. Complex problem-solving: a field in search of a definition? Theoretical Issues in Ergonomics Science, 6 (1), 5–33. doi.10.1080/14639220512331311553

Righi, A., Saurin, T.A., 2015. Complex socio-technical systems: characterization and management guidelines. Applied Ergonomics, 50, 19–30.

Rioux, F., Bernier, F., Laurendeau, D., 2008. Multichronia–A Generic Visual Interactive Simulation Exploration Framework. In Proceedings of the Interservice/Industry Training, Simulation and Education Conference.

Rouse, W.B., Morris, N.M., 1986. On looking into the black box: Prospects and limits in the search for mental models. Psychological bulletin, 100 (3), 349. doi:10.1037//0033-2909.100.3.349

Scaife, M., Rogers, Y., 1996. External cognition: how do graphical representations work? International journal of human-computer studies, 45 (2): 185-213. doi:10.1006/ijhc.1996.0048

Shah, A.K., Oppenheimer, D.M., 2008. Heuristics made easy: an effort-reduction framework. Psychological bulletin, 134 (2), 207. doi. 10.1037/0033-2909.134.2.207

Shneiderman, B., 2008. Extreme visualization: squeezing a billion records into a million pixels. In Proceedings of the 2008 ACM SIGMOD international conference on Management of data (pp. 3–12). ACM

Slof, B., Erkens, G., Kirschner, P.A., Jaspers, J.G.M., Janssen, J., 2010. Guiding students' online complex learning-task behavior through representational scripting. Computers in Human Behavior, 26 (5), 927–939. doi. 10.1016/j.chb.2010.02.007

Spector, J.M., 2000. System dynamics and interactive learning environments: Lessons learned and implications for the future. Simulation & Gaming, 31 (4), 528–535. doi:10.1177/104687810003100406

Stekler, H.O. 2007. The future of macroeconomic forecasting: Understanding the forecasting process. International Journal of Forecasting, 23 (2), 237–248. doi:10.1016/j.ijforecast.2007.01.002

Sterman, J.D., 1994. Learning in and about complex systems. System Dynamics Review, 10 (2-3), 291-330. doi. 10.1002/sdr.4260100214

Sterman, J.D., 2000. Business dynamics: Systems thinking and modeling for a complex world. Irwin/McGraw-Hill.

Sterman, J.D., 2006. Learning from evidence in a complex world. American journal of public health, 96 (3), 505–14. doi:10.2105/AJPH.2005.066043

Stull, A.T., Mayer, R.E., 2007. Learning by doing versus learning by viewing: Three experimental comparisons of learner-generated versus author-provided graphic organizers. Journal of educational psychology, 99 (4), 808–820. doi. 10.1037/0022-0663.99.4.808

Surdu, J.R., Kittka, K., 2008. Deep Green: Commander's tool for COA's Concept. Proceedings of the 2008 Computing, Communications and Control Technology Conference (DDT) 29 Jun – 2 Jul, Orlando Fl.

Thomas, J.J., Cook, K.A., eds., 2005. Illuminating the path: The research and development agenda for visual analytics. IEEE Computer Society Press.

Thomas, J.J., Cook, K.A., 2006. A visual analytics agenda. IEEE Computer Graphics and Applications, 26 (1), 10–13. doi. 10.1109/MCG.2006.5

Tominski, C., Schulze-Wollgast, P., Schumann, H., 2008. Visual methods for analyzing human health data. Encyclopedia of healthcare information systems, 1357–1364. doi. 10.4018/978-1-59904-889-5.ch170

Trowbridge, J.E., Wandersee, J.H., 1994. Identifying critical junctures in learning in a college course on evolution. Journal of Research in Science Teaching, 31 (5), 459-473. doi. 10.1002/tea.3660310504

Tye-Gingras, O., 2011. Intégration et exploitation d’outils de visualisation - Compréhension de situations complexes. Mémoire de maîtrise, Université Laval, Québec, Canada

Van Dam, A., Forsberg, A., Laidlaw, D.H., LaViola, J.J., Simpson, R.M., 2000. Immersive VR for scientific visualization: A progress report. IEEE Computer Graphics and Applications, 20 (6): 26–52. doi. 10.1109/38.888006

Vester, F., 2007. The art of interconnected thinking: Ideas and tools for tackling with complexity. Munich: MCB-Verlag.

Waern, Y., Cañas, J.J., 2003. Microworld task environments for conducting research on command and control. Cognition, Technology & Work, 5 (3), 181–182. doi:10.1007/s10111-003-0126-y

Yasarcan, H., 2010. Improving understanding, learning, and performances of novices in dynamic managerial simulation games. Complexity, 15 (4), 31–42. doi. 10.1002/cplx.20292

Figure 1. The interactive simulation was supported by Multichronia, a generic visual interactive simulation exploration framework providing users with a visual history of a series of simulation. Multichronia supports hypothesis-testing by allowing the creation of multichronic trees which let users simulate an instance and observe what would happen if a particular parameter was changed at any given moment in time. Whenever a parameter is changed, a new “branch” is created which visually evolves alongside its originating branch, thus making comparison between the original and the newly created instance easy for users. Branches correspond to what-if scenarios that may help understand the effects of specific changes through time.

Figure 2. Analytical interactive visualizations have been developed with the Eye-Sys software package (http://www.eye-sys.com). Eye-Sys allows the gathering and manipulation of data in real-time from different sources for driving interactive, real-time visualizations. Within IMAGE, Eye-Sys provided a scientific dashboard with a variety of views aimed at facilitating the comparison of outcomes as a function of the parameter allocations. A ribbon view allowed the participant to inspect the development of MoPs across 100 iterations (top left). A cubic view allowed one to assess the impact of any three variables for each simulation (top right). The main difference between Eye-Sys in IMAGE-Desktop and IMAGE-CAVE was the presence of immersive views for the latter, whereas the former only offered traditional interactive views. Finally, a tactical view allowed analysts to see actual convoy movements (bottom).

Figure 3. CoGUI-IMAGE is a tool for editing conceptual graphs. It was built by Defence R&D Canada (DRDC) from an existing tool called CoGUI, developed by the Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier.

Figure 4. Experimental setup in the IMAGE-Desktop condition. The screens were disposed in arc in front of the participant. Each screen was assigned to a specific function, namely knowledge representation (left), interactive simulation (center) and enhanced visualization (right). Although not described in the present article, participants were required to wear a

head-tracking device which was used for behavioural measurement and design recommendations.

Figure 5. Experimental setup for the IMAGE-CAVE condition. The CAVE operated in two distinct modes: immersive and non-immersive display. In the immersive mode, participants stood and moved around the workspace at will. An extra input device, the WAND, was added to the IMAGE toolkit in order to reproduce mouse functionalities while standing. The WAND consisted of a thumb-stick for navigating and five programmable buttons. Participants in the IMAGE-CAVE condition may also have used electronic shutter glasses (stereo-glasses) to explore two immersive interactive data visualizations called the cube view (lower left panel) and the ribbon view (lower right panel).

Figure 6. The structural knowledge test consisted of a matrix with variable names in rows and columns, and cells allowing the participant to estimate the linear correlation (between -1 and +1) between the variables of each row and column.

Figure 7. Decision making performance by experimental condition. Error bars represent mean squared error.

Figure 8. Structural knowledge by experimental condition. Error bars represent mean squared error.