Defence Research and Development Canada
Recherche et développement pour la défense Canada

DEFENCE R&D DÉFENSE

IMAGE Final Report

An Interactive Computer-Aided Cognition Capability

Michel Lizotte
François Bernier
Marielle Mokhtari
Éric Boivin
DRDC Valcartier

Defence Research and Development Canada – Valcartier

Technical Report
DRDC Valcartier TR 2012-397
December 2012

Canada

IMAGE Final Report

An Interactive Computer-Aided Cognition Capability

Michel Lizotte
François Bernier
Marielle Mokhtari
Éric Boivin
DRDC Valcartier

Defence Research and Development Canada – Valcartier

Technical Report
DRDC Valcartier TR 2012-397
December 2012

IMPORTANT INFORMATIVE STATEMENTS

Abstract

With the advances in Science and Technology (S&T), we are now able to build artefacts whose complexity approaches those of life itself. To address the new defence and national security challenges of dealing with an ever increasing range of possibilities – some natural but many human-made – DRDC scientists propose a toolset concept that enables the collaboration of experts from different disciplines so that scientists can try to reach a shared understanding of a complex situation. The toolset concept supports the need for disentangling complex situations more quickly by using the synergy between knowledge representation, modeling, simulation, and visualisation technologies. Many Canadian Forces contexts could potentially benefit from such a toolset in order to accelerate understanding. A whole-of-government team working out a situation involving security, development, and political issues is an obvious context, but facing a cyber-threat requires an efficient collaboration between experts in order to analyze friendly plans, enemy intents, computer networks, and software applications.

Résumé

Avec les progrès en sciences et technologie (S&T), l'homme est désormais capable de construire des objets dont la complexité se rapproche de celle de la vie elle-même. Pour aborder les nouveaux défis de sécurité nationale et de défense devant traiter un éventail toujours croissant d'éléments, certains naturels, mais beaucoup fabriqués par l'homme, les scientifiques de RDDC proposent un concept de boîte à outils visant la collaboration d'experts œuvrant dans différentes disciplines qui tentent de parvenir à une compréhension commune d'une situation complexe. Ce concept vise à soutenir le besoin de déchiffrer des situations complexes plus rapidement en utilisant la synergie entre la représentation des connaissances, la modélisation, la simulation et les technologies de visualisation. Les Forces canadiennes pourraient potentiellement bénéficier de ces outils pour accélérer la compréhension de nombreux contextes dans lesquels elles doivent être déployées. Une équipe pangouvernementale travaillant sur une situation impliquant des questions politiques, de sécurité et de développement est un contexte évident. Par contre, faire face à une menace cybernétique est un contexte nécessitant une collaboration efficace entre les experts de la planification côté ami, les experts des intentions de l'ennemi, ainsi que ceux des réseaux informatiques et des applications logicielles.

This page intentionally left blank.

Executive summary

IMAGE Final Report: An Interactive Computer-Aided Cognition Capability

Michel Lizotte; François Bernier; Marielle Mokhtari; Éric Boivin; DRDC Valcartier TR 2012-397; Defence R&D Canada – Valcartier; December 2012.

Complexity has always existed in natural and biological realms. However, with the advances in Sciences and Technology (S&T), humanity is now capable of building artefacts whose complexity approaches those of life itself. There is a need for using new methods to tackle this new complexity. To address the new defence and national security challenges of dealing with an ever increasing range of possibilities – some natural but many human-made – DRDC scientists propose a toolset concept based on the collaboration of experts from different disciplines trying to reach a shared understanding of a complex situation (CS). This concept aims at supporting the need for disentangling CSs more quickly.

The IMAGE project was part of the Technological Investment Funds (TIF) program. Researchers from various research fields and organizations contributed to the effort: software, computer, and electrical engineering at DRDC Valcartier and Université Laval (Computer Vision and Systems Laboratory); cognitive psychology at Université Laval (Cognition Distribution Organisation Technologies Laboratory); and operational research at DRDC Centre for Operational Research and Analysis.

The concept IMAGE: The proposed concept aims at disentangling CSs more quickly, specifically targeting the collaboration of experts from different disciplines trying to reach a shared understanding of a CS. The goal of the concept is two-fold: increasing understanding of a CS, and enabling individuals to share their understanding. The concept elaboration was mainly guided by three principles or assumptions: (1) Iterative understanding: a common understanding reached through revision and sharing of successive representations of the situation; (2) Human-driven toolset: the common understanding is above all a human task involving team members using tools in their own way; and (3) Synergy of technologies: the common understanding is better served by the synergy of technologies than a single one. IMAGE researchers have identified four complementary modules to achieve this synergy: Representation, Scenarisation, Simulation, and Exploration. The Representation module (I-REP) consists in transitioning a mental model into a comprehension model. From this comprehension model, the Scenarisation module (I-SCE) generates an executable model to be used by the Simulation module (I-SIM), which provides a high level of interactivity and enables the investigation of the space of variables. Finally the Exploration module (I-EXP) provides powerful tools allowing experts to invent views that bring a significant meaning to large datasets generated by the I-SIM module.

Prototypes: In order to validate the usefulness and the feasibility of the concept, the IMAGE team built two prototypes: IMAGE Version 1 (V1) and IMAGE Version 2 (V2). The first prototype was a partial implementation of three (3) modules, Representation, Simulation, and Exploration, to be used during the cognitive experimentation. The Scenarisation module was not required. Testing the concept IMAGE is a very audacious task because usability and training are critical elements and need to be realised with immature research tools. IMAGE experimentation

was not only conducted to test the concept and bring evidence (with which the imagined toolset could augment the comprehension of CSs) but also to improve its implementation. Operational research scientists brought know-how about military situations and contributed to validate the solution and the experimental approach. The second prototype demonstrated the feasibility of more advanced functionality, including collaboration, immersive capabilities, and a scenarisation module. Both prototypes used a small-scale CS on the theme of convoy.

Experimentation: Fifty-eight people have undergone between nine (9) and fourteen hours (depending on the group they were in) of effort distributed over several sessions. This experimentation addressed three questions: does IMAGE improve comprehension of CS; what is the added value of immersion for IMAGE; what are the typical and the efficient ways to exploit IMAGE.

Four measures were selected to help answer the first two questions. The two performance metrics, Analysis and Prediction, were evaluated using questions. Participants had tool access during analysis efforts while they did not have tool access during prediction efforts. The Quality of mental model measured how well participants could estimate correlations between every combination of the most important variables. Finally, the Response selection fallacy (RSF) provided the “error-reduction” potential of such a toolset. During analysis sessions, a participant explored the parameter space by instantiating simulations and then picked one simulation in order to answer a question. The RSF measure corresponds to the proportion of instantiated simulations better answering a question.

In the vast majority of measures taken, on the average, IMAGE conditions participants outperformed those of the Baseline condition; however, many of these results were not statically significant due to high individual variability in this kind of experiment. Nevertheless, experimental results demonstrated that IMAGE V1 contributes to improving analysis performance, especially when dealing with manageable level of complexity (called low level of complexity in the experiment). Improved tools would likely increase comprehension at any level of complexity because IMAGE group participants performed better than Baseline in all measures even though these were not statistically significant. The Prediction performance and Quality of mental model metrics both showed that IMAGE provides a benefit but only at the lowest level of complexity, suggesting that some situations cannot be handled by humans without the calculation power and memory accuracy of a computer: the situations are just too complex. Analyses of the RSF metric indicated that about half of the benefit of IMAGE over Baseline was attributable to the ability of IMAGE to increase comprehension; the remaining being explained by the error reduction potential of IMAGE.

No difference on the measures of comprehension between the IMAGE-CAVE and IMAGE-Desktop conditions were statistically significant; but some reasons may explain the lack of immersion benefits. First, usability of I-SIM in immersive mode was inferior mainly because of its lack of responsiveness and its implementation on a tablet PC. In addition, IMAGE V1 in immersive mode did not offer a full user-centric interactive experience, which is likely a key interface in order to benefit from the approach. However, such a user-centric approach would have been too different from the IMAGE Desktop condition for an adequate comparison within a cognitive experimentation setup.

Various means were used to examine how people are using IMAGE modules. From an I-REP perspective, the results suggest that the elaboration of a comprehension model has no plus-value once the problem is well defined. However, the automatic or assisted generation of graphs is a relevant approach, and Representation artefacts usefulness increase with the complexity of the situation. Finally, the CAVE wider screen seems to improve work intensity on graphs but the results are not statistically meaningful. I-SIM was the most used tool and the only one for which use time was positively correlated with performance. The parameter space exploration strategies supported by I-SIM were similar between the IMAGE and Baseline conditions, suggesting that I-SIM does not interfere significantly with the model study process. The study of results from an I-EXP perspective demonstrates an obvious intertwining time sharing usage with I-SIM. Even if the I-EXP views were not rigorously introduced to participant (especially for the desktop tutorial), and therefore were underused, I-EXP focus time in percentage reaches quarter (SOP) to half (Tactical) of session time. Unfortunately, correlation with performance was not demonstrable because experimental design questions did not focus on I-EXP capabilities. Some views functionality stayed useless despite their advanced analysis capability. Few participant comments point in this direction.

An issue that would have to be addressed in the future is the lack of statically significant results due to individual variability in this kind of experiment. As mentioned, IMAGE conditions participants performed better than Baseline ones on the average, but very often the results were not significant due to individual variability.

Potential applications: There are many contexts into which the IMAGE concept has potential. Two different examples, both about circumstances where there is a need to disentangle CSs more quickly, are: a comprehensive approach, and a cyber-threat. The first one is operational while the second one has more a technological flavour, although the second example may have important operational impacts.

Canadian Forces (CF) recognises the importance of a comprehensive approach that considers all dimensions of a situation, over and above that of military power alone. The various issues (ethnics, religious, ideological, and material), the various power and influence (diplomatic, economic, informational ...) as well as the national and international public opinion and Medias are examples of the variety and diversity of the dimensions. A comprehensive approach is required to elaborate plans but, most important, a whole-of-government approach is required to achieve the strategic national objectives.

Cyber is a quite new and complex challenge. Actually, CF created the new environment Cyber, distinct from Land, Air, and Sea, to face that challenge. Risk analysis identifying critical components, investigating attack vectors, and studing impacts of successful attacks, is needed to face cyber threats impacting freedom of CF manoeuvre. Such analysis requires an efficient collaboration between experts of friendly plans, enemy intents, computer networks, and software applications.

Sommaire

IMAGE Final Report : An Interactive Computer-Aided Cognition Capability

Michel Lizotte; François Bernier; Marielle Mokhtari; Éric Boivin; DRDC Valcartier TR 2012-397 ; R & D pour la défense Canada – Valcartier ; décembre 2012.

La complexité a toujours existé dans les mondes naturels et biologiques. Cependant, avec les progrès en sciences et technologie (S&T), l'homme est désormais capable de construire des objets dont la complexité se rapproche de celle de la vie elle-même. De nouvelles méthodes sont nécessaires pour s'attaquer à cette nouvelle complexité incluant de tels objets. Pour aborder les nouveaux défis de sécurité nationale et de défense devant traiter un éventail toujours croissant d'éléments, certains naturels, mais beaucoup fabriqués par l'homme, les scientifiques de RDDC proposent un concept de boîte à outils visant la collaboration d'experts œuvrant dans différentes disciplines qui tentent de parvenir à une compréhension commune d'une situation complexe (SC). Ce concept vise à soutenir le besoin de déchiffrer des SCs plus rapidement.

Le projet IMAGE faisait partie du programme de Fonds d'investissement technologique (FIT). Des chercheurs provenant de domaines de recherche et d'organisations différents ont contribué à l'effort: logiciel, génie électrique et génie informatique à RDDC Valcartier et à l'Université Laval (Laboratoire de vision et systèmes numériques); psychologie cognitive à l'Université Laval (Laboratoire Cognition – Distribution – Organisation – Technologies), et recherche opérationnelle au Centre d'analyse et de recherche opérationnelle de RDDC.

Le concept IMAGE : Le concept proposé vise à déchiffrer les SCs plus rapidement, ciblant spécifiquement la collaboration d'experts de différentes disciplines qui tentent de parvenir à une compréhension commune d'une SC. L'objectif du concept IMAGE est double: accroître la compréhension d'une SC et permettre aux individus de partager leur compréhension. L'élaboration du concept a principalement été guidée par trois principes ou hypothèses: (1) la compréhension itérative: une compréhension commune est atteinte grâce à la révision et au partage de représentations successives de la situation, (2) le contrôle des outils par l'homme: la compréhension commune est avant tout une tâche réalisée par l'homme et qui implique des membres de l'équipe utilisant les outils à leur manière, et (3) la synergie des technologies: la compréhension commune est mieux servie par la synergie de technologies que par une seule. Les chercheurs de IMAGE ont identifié quatre modules complémentaires pour atteindre cette synergie: Représentation, Scénarisation, Simulation et Exploration. Le module de Représentation (I-REP) consiste à expliciter un modèle mental en un modèle de compréhension. À partir de ce modèle de compréhension, le module de Scénarisation (I-SCE) génère un modèle exécutable utilisé par le module de Simulation (I-SIM) qui apporte un niveau élevé d'interactivité permettant une investigation de l'espace des variables. Enfin, le module d'Exploration (I-EXP) fournit des outils puissants permettant aux experts d'inventer des vues qui apportent une signification importante aux grands ensembles de données générés par le module I-SIM.

Prototypes : Afin de valider l'utilité et la faisabilité du concept, l'équipe IMAGE a construit deux prototypes: IMAGE Version 1 (V1) et IMAGE Version 2 (V2). Le premier prototype est

une implémentation partielle de trois (3) modules, Représentation, Simulation et Exploration, utilisés lors de l'expérimentation cognitive. Le module de Scénarisation n'était pas nécessaire pour l'expérimentation. Tester le concept IMAGE est une tâche audacieuse car l'utilisabilité et la formation sont des éléments essentiels et doivent être réalisés avec des outils de recherche immatures. L'expérimentation IMAGE a été menée non seulement pour tester le concept et apporter la preuve que la boîte à outils imaginée peut augmenter la compréhension des SCs, mais aussi pour améliorer sa mise en œuvre. Un scientifique de recherche opérationnelle a apporté ses connaissances quant aux situations militaires et a contribué à valider la solution et l'approche expérimentale. Le second prototype a démontré la faisabilité de fonctionnalités plus avancées dont la collaboration, les capacités immersives et le module de Scénarisation. Les deux prototypes ont utilisé une SC à petite échelle sur le thème du convoi militaire.

Expérimentation : Cinquante-huit personnes ont fourni entre neuf (9) et quatorze heures (selon le groupe auquel ils appartenaient soit la condition de base, soit les conditions IMAGE) d'efforts réparties sur plusieurs séances. Cette expérimentation posait trois questions. Est-ce que IMAGE améliore la compréhension des SCs? Quelle est la valeur ajoutée de l'immersion pour IMAGE? Quels sont les moyens typiques et efficaces d'exploiter IMAGE?

Quatre mesures ont été sélectionnées pour aider à répondre aux deux premières questions. Les indicateurs de performance d'*Analyse* et de *Prédiction* ont été évalués à l'aide de questions. Les participants ont eu accès aux outils durant la session d'*Analyse* alors que non durant la session de *Prédiction*. La *Qualité du modèle mental* a mesuré à quel point les participants ont pu estimer les corrélations entre toutes les combinaisons des variables les plus importantes. Enfin, l'*Erreur de sélection de la réponse* a fourni le potentiel de « réduction d'erreur » d'un tel ensemble d'outils. Durant les sessions d'*Analyse*, le participant a exploré l'espace des paramètres en instanciant des simulations et en choisissant une simulation pour répondre à une question. La mesure de l'*Erreur de sélection de la réponse* correspond à la proportion de simulations instanciées répondant le mieux à une question.

Pour la grande majorité des mesures prises, les conditions IMAGE ont surpassé la condition de base sur la moyenne, mais beaucoup de mesures n'étaient pas statistiquement significatives en raison de la variabilité individuelle élevée dans ce type d'expérience. Néanmoins, les résultats expérimentaux montrent que IMAGE V1 contribue à améliorer les performances d'*Analyse*, surtout lorsqu'il s'agit d'un niveau gérable de complexité (appelé faible niveau de complexité dans l'expérimentation). L'amélioration des outils augmenterait probablement la compréhension à tous les niveaux de complexité puisque les participants aux conditions IMAGE ont systématiquement mieux performé que ceux évoluant dans la condition de base, mais les résultats n'étaient pas statistiquement significatifs. Les métriques de performance de *Prédiction* et de *Qualité du modèle mental* montrent toutes deux que IMAGE fournit un avantage, mais seulement au niveau le plus bas de la complexité, ce qui suggère que certaines situations ne peuvent être manipulées par l'homme sans la puissance de calcul et la précision de la mémoire d'un ordinateur: elles sont tout simplement trop complexes. Les analyses de la métrique de l'*Erreur de sélection de la réponse* indiquent que près de la moitié du bénéfice des conditions IMAGE par rapport à la condition de base était attribuable à la capacité de IMAGE à accroître la compréhension; le reste étant expliqué par le potentiel de réduction d'erreur de IMAGE.

Aucune différence sur les mesures de la compréhension entre les conditions IMAGE-CAVE et IMAGE-Bureau n'est statistiquement significative, mais certaines raisons peuvent expliquer l'absence de bénéfice de l'immersion. Tout d'abord, l'utilisabilité de l'outil I-SIM dans le mode immersif était médiocre principalement en raison de son manque de réactivité et de sa mise en œuvre sur une tablette PC. En outre, IMAGE V1 dans le mode immersif n'offrait pas une expérience interactive complète centrée sur l'utilisateur, ce qui est sans aucun doute la clé pour obtenir des bénéfices. Cependant, une approche centrée sur l'utilisateur aurait été trop différente de la condition IMAGE-Bureau pour une comparaison adéquate dans une configuration d'expérimentation cognitive.

Divers moyens ont été utilisés pour examiner la façon dont les gens utilisent les modules IMAGE. Du point de vue de I-REP, les résultats suggèrent que l'élaboration d'un modèle de compréhension n'a pas de plus-value une fois que le problème est bien défini. Cependant, la génération automatique ou assistée des graphes est une approche pertinente et l'utilité des objets de Représentation augmente avec la complexité de la situation. Enfin, le grand écran du CAVE semble améliorer l'intensité du travail sur les graphes, mais les résultats ne sont pas statistiquement significatifs. L'outil I-SIM est le plus utilisé, et le seul pour lequel le temps d'utilisation était positivement corrélé avec la performance. Les stratégies d'exploration de l'espace des paramètres soutenues par I-SIM étaient similaires entre la condition de base et les conditions IMAGE, ce qui suggère que I-SIM n'interfère pas de manière significative avec le processus d'étude du modèle. L'étude des résultats du point de vue de I-EXP démontre un évident entrelacement en temps partagé avec I-SIM. Même si les vues de I-EXP ne furent pas rigoureusement présentées aux participants (en particulier pour le tutoriel de IMAGE-Bureau), et ont donc été sous-utilisées, le temps d'exploitation de I-EXP en pourcentage se situe entre le tiers (SOP) et la moitié (tactique) du temps de session. Malheureusement, la corrélation avec la performance n'est pas démontrable car les questions posées lors de l'expérimentation ne se concentrent pas sur les capacités de I-EXP. Certaines fonctionnalités de vues sont restées sans usage en dépit de leur capacité d'analyse avancée. Quelques participants commentent dans cette direction.

Une question qui devrait être abordée dans les travaux futurs est le manque de résultats statistiquement significatifs dus à la variabilité individuelle dans ce type d'expérimentation. Comme mentionné précédemment, les participants évoluant sous les conditions IMAGE étaient meilleurs que ceux évoluant sous la condition de base sur la moyenne, mais très souvent, les résultats n'étaient pas significatifs en raison de la variabilité individuelle.

Applications potentielles: Il existe de nombreux contextes pour lesquels le concept IMAGE a un potentiel. Voici deux exemples où il est nécessaire de déchiffrer plus rapidement les SCs: une approche globale et une menace cybernétique. Le premier est un exemple opérationnel tandis que le second est un exemple à saveur plus technologique, même s'il peut avoir d'importantes répercussions opérationnelles.

Les Forces canadiennes (FC) reconnaissent l'importance d'une approche globale tenant compte de toutes les dimensions d'une situation, plus que la puissance militaire seule. Les diverses questions (d'ordre ethnique, religieuse, idéologique et matérielle), la puissance et l'influence (d'ordre diplomatique, économique, informationnel ...) ainsi que les opinions publiques nationale et internationale et les médias sont des exemples de la variété et de la diversité de dimensions. Une approche globale est nécessaire pour élaborer des plans mais une approche pangouvernementale est plus que nécessaire pour atteindre les objectifs nationaux stratégiques.

La cybernétique est un défi tout à fait nouveau et complexe. En fait, les FC ont créé le nouvel environnement Cyber, distinct de Terre, Air et Mer, pour y faire face. L’analyse de risques identifiant les composants critiques, investiguant les vecteurs d’attaque et les impacts de l’étude des attaques réussies est nécessaire pour faire face aux menaces cybernétiques ayant une incidence sur la liberté de manœuvre des FC. Une telle analyse nécessite une collaboration efficace entre les experts de la planification côté ami, ceux des intentions ennemies, ainsi que les experts des réseaux informatiques et des applications logicielles.

This page intentionally left blank.

Table of contents

List of figures

List of tables

Acknowledgements

The authors would like to thank Dr. Denis Poussart for his important contribution during the inception phase of the project and his precious advice without whom the project would not have been possible.

We would also want to highlight the important contributions of Frédéric Drolet, Claire Lalancette and Mustapha Zaraket from the System of Systems section for helping to manage, design and implement the different phases of IMAGE.

We would also like to thank the following people from Université Laval who contributed significantly to the project with their time, their intellectual input or their devotion:

- from the Cognitive Psychology Research Group, School of Psychology : Dr Sébastien Tremblay (professor), Dr Patrick Jeuniaux, Dr Daniel Lafond (postdoctoral researchers), Jean-François Gagnon, Bernard Lebel, Julie Champagne and Louis Dumont (graduate students);
- from the Computer Vision and Systems Laboratory, Department of Electrical and Computer Engineering: Dr Denis Laurendeau (professor), Dr Sylvain Comtois, Denis Ouellet (research professionals), Maxime Girardin, Julien-Charles Lévesque, Étienne Ouellet, François-Michel de Rainville, François Rioux and Olivier Tye-Gingras (graduate students); and
- from the Department of Computer Science and Software Engineering: Maxime Pereira-Morneau and Hugo St-Louis.

We also wish to thank other contributors to this project both from DRDC Valcartier, namely Dr Michel B. DuCharme, and from DRDC CORA, Paul Massel and Dr Irene Pestov.

Finally, we wish to extend our gratitude to the people, most from DRDC Valcartier and Université Laval, who gave of their time to participate in the studies described in this report.

1 Introduction

Complexity has always existed in natural and biological realms. However, with the advances in Sciences and Technology (S&T), humanity is now capable of building artefacts whose complexity approaches those of life itself. There is now a need for using new methods to tackle this new complexity including such artefacts. There are tremendous increases in speed, density, spatial scope, and coupling between an ever increasing range of elements, some natural, many synthetic. The human mindset, however, is still deeply rooted in the classical concept of reductionism, in which a problem is solved by decomposing it into sub-problems.

Complexity and the military

Complexity of situations is a fact that the military have to deal with, a core challenge of the new defence landscape. In addition to the rapid technology evolution providing more information and capabilities combined with challenging terrains, e.g., urban environments, the "human terrain" involves insurgents, mixed populations, non-government organizations, failed states, and so on. This last leads to a need for joint and / or multinational coordination and synchronization.

Nevertheless these complex situations (CSs), military still, now more than ever require timely decisions to overmatch the threat. In addition, decision quality will always be closely associated with the level of the situation comprehension. Nowadays, understanding is quite a challenge, with the rapid operational and technological changes and key aspects that arise from interactions between multiple, diverse, interconnected elements bringing a strong need for better and faster disentangling CSs. Approaching complexity in an effective way requires the capability of teaming people with different backgrounds because such problems cannot be studied in isolation with a reductionism approach.

The research project

To address this problem, a Technology Investment Fund (TIF) project called IMAGE has been carried out by a team of scientists from Defence Research and Development Canada (DRDC) and Université Laval between April 2006 and March 2011. The purpose was to study CSs and imagining a toolset that could help the Human to better deal with them. The current report summarises the principal findings of the project. The reader is invited to consult previously published reports and papers for more details. These publications are cited in due course in the report.

Researchers from various research fields contributed to the effort. From an engineering point of view, the objective was to conceive and demonstrate the feasibility of a set of software tools, using knowledge representation, simulation, visualization and immersive technologies, to assist a team of experts in developing a better understanding of CSs. In terms of cognitive psychology, the objective was to bring evidence that the imagined toolset can augment the comprehension of CSs. Finally, operational research brought know-how about military situations and contributed to validate the IMAGE solution and experimental approach, including a realistic definition of a Baseline toolset to be used as the control condition during the cognitive experimentation.

The current document summarises the IMAGE project results that can be split as follow: the overall concept itself; modules composing the concept along with prototypes implementing some functionality and; results of the cognitive experimentation.

Concept IMAGE

Chapter 2 provides the rationales behind the effort and introduces the toolset concept. This concept called IMAGE aims at increasing understanding of a CS and enabling individuals to share their comprehension and exploits knowledge representation, simulation and visualization technologies.

Chapter 3 introduces two CSs about land convoys facing insurgents constituting the case study used for this research. These CSs are utilized to develop IMAGE modules, experiment and test their feasibility of some advanced functionality.

IMAGE modules

Four (4) chapters are describing the modules composing the concept: Representation, Scenarization, Simulation and Exploration.

The Representation module, summarized in Chapter 4, allows expert to explicit their comprehension using conceptual graphs. Representation is closely related to the mental model that human are gradually forming to understand their environment. Drawings are often used to better assimilate a problem or share it with others; this is Representation is about, expliciting knowledge to better understand.

The Scenarization module, introduced in Chapter 5, assists experts in the creation of an executable model from the comprehension graphs created with the Representation module. As for the Representation module, it aims at reproducing techniques often used by people but with the computer power. To better approach the future, a simulation based on current beliefs about a situation is frequently used by human. It can be either a mental act, such as a visualisation technique, or an actual exercise as the military do war-gaming to fake dynamics between friendly and adversary forces.

Chapter 6 describes the Simulation module that actually exploits the executable model resulting from Scenarization. Computer simulation allows a person to confront preconceptions and explore a CS thoroughly in a systematic and repetitive manner that should help improve comprehension. The Simulation module controls the parameters values, execution, and branching in order to assist an expert performing what-if analysis.

Chapter 7 presents the Exploration module that enable experts to make sense from large datasets generated by the various simulations. To achieve this, the module combines visualization and rich user interactions to explore, organize and process information. The module consists in making datasets clear in tailored views, which can be used by experts to augment their individual and collective comprehension.

Cognitive experimentation and findings

The experimental results and their interpretation are described in Chapter 8. The experiment was designed to assess how the human comprehension capacities would be augmented with IMAGE suite of tools in comparison with usual tools. This chapter presents a summary of the method and results but a complete report on IMAGE experiment (Gagnon, Jeuniaux, Lafond, & Tremblay, 2012) contains more details on many aspects presented here. In addition, the chapter provides many additional results and interpretations that are not included in the experiment report.

Finally, Chapter 9 summarizes the main findings and proposes potential applications for the concept IMAGE.

This page intentionally left blank.

2 Concept IMAGE: Disentangling complex situations

The world is complex: financial markets, epidemic diseases, software-intensive systems, military capabilities, theatres of operations. One must understand to decide and deal with such situations: public investments, vaccination campaigns, software engineering, capability engineering, military operations... More than ever, the issue is about timely decisions to overmatch the threat. Such contexts require decisions before well understanding a situation although more one knows about it, better are chances of success. This leads to a strong need for better and / or faster to disentangle CSs. What are the characteristics of CSs? How could software tools assist the Human to better deal with CSs? These are the main questions address by the current chapter.

2.1 A fact: Complex situations

Complexity of situations is a fact that the military have to deal with, a core challenge of the new defence landscape: operations involve far more diversified surroundings and a wider range of ways and means, all acting with renewed synergy. The current and anticipated types of operations are characterized by a high level of complexity. There is a need for joint and / or multinational coordination and synchronization. The “human terrain” involves insurgents, mixed populations, non-government organizations, failed states… The physical terrain is not easy e.g. urban environments. In addition technologies are evolving faster than ever and provide more information and capabilities. These are some examples of CS elements.

Many of these elements are hard to capture and assess. They are highly context-dependant and consequently bring ontology challenges. They occur at multiple levels, over a wide spectrum, from critical real-time to long term and over varying domains, some physical, and some non-tangible. As a consequence, current operational and technological conditions lead to situations in which the quality and timeliness of the decision process can be swamped in a sea of complexity. Acute concerns in contemporary defence and national security are characterized by the interweaving of multiple and diverse elements, leading to global properties that emerge from the dynamics of interactions.

There is no single definition of complexity but for the purpose of the current work, it is described from three perspectives: structural, behavioural and consequence. From a structural perspective, a CS has usually many linked, entangled elements, multiple scales in time, size and resolution, no centralized command and unclear boundaries. From a behavioural perspective, a CS involves systems that may communicate through their environment (stigmergy), self-organize (adapt), reach multiple equilibrium points (attractors), which evolution depends on their life history and may become counter-intuitive. As a consequence, systems evolving in such situations have to be resilient, operate at the edge-of-chaos but are hard to predict and study in isolation.

A CS must not be confounded with a complicated one. A complicated situation may comprise a high number components and/or a very high number of variables, but its behaviour can be predicted by experts and appropriate tools. For instance, the hardware of a computer is complicated in the sense that it comprises a very high number of interacting processors, transistors, and buses that few experts understand. Nevertheless, it behaves in a predicted way. At the opposite, some logistic equations (one-dimensional feedback system with only one variable and one control parameter) that model complex real-world phenomena such as long-term change in a species population look simple. They are nevertheless complex in the sense they are highly sensitive to the value of the control variable and, if a natural phenomenon behaves according to these equations, predicting their future states is hardly feasible.

Many situations, when they become more complicated, are more likely to include complex elements that will eventually “contaminate” other parts, making the whole situations complex. Most real-word CSs seem also complicated for this reason, but often the complexity parts are not the same as those making it complicated. In the IMAGE project however, while complexity was a must, the complicated aspect was considered a nuisance for an experiment perspective because it would have interfered with and hidden the complexity during the little amount of time available to participants. As a consequence, the reader must not underestimate the level of complexity of the CS used for the experiment based on the fact they seem non-complicated at first glance.

2.2 A requisite: Collaboration of experts

Good decisions usually results from a good comprehension. Nowadays, such understanding is quite a challenge with rapid operational and technological changes. This is a new challenge where key behaviours and qualities arise from interactions between multiple, diverse, interconnected elements. Approaching such complexity in an effective way require the capability of teaming people with different backgrounds – working in a collaborative manner to effectively explore scenarios and problem spaces with multiple dimensions. These dimensions have fuzzy boundaries, are abstract and difficult to assess and cannot be studied in isolation.

For instance, engineering a military capability necessitates a close collaboration between various actors, experts from different disciplines who need to work out a solution considering all factors such as personnel, doctrine, infrastructure and equipment. Moreover each of these factors has its own complexity and may involve different specialists of the discipline. The preceding factors are more related to the performance of the future capability. In addition, risks, costs and schedule may require participation of other kind of specialists. Better is the collaboration, better will be the solution which will likely involve many trade-offs between the various facets of the problem whether related to the performance (e.g. sensors, protection and fire power for a vehicle capability) or not (delivery date and cost of the vehicle).

Operationally, Canadian Forces (CF) also recognises the importance of a comprehensive approach considering all dimensions of a situation, more than the military power. The various issues (ethnics, religious, ideological and material), the various power and influence (diplomatic, economic, informational...) as well as the national and international public opinion and Medias are examples of the variety and diversity of the dimensions. A comprehensive approach is required to elaborate plans but moreover, a whole-of-government approach is required to achieve the strategic national objectives.

For instance, Canada provided a support to Haiti through the work of many Canadian departments and agencies working from the Embassy in Port-au-Prince: Foreign Affairs and International Trade Canada, the Canadian International Development Agency, Public Safety, the Royal Canadian Mounted Police, the Correctional Service of Canada, Citizenship and Immigration Canada and the Department of National Defence. Each department and agency played a key role in delivering Canadian support to Haiti. Another example is the Canadian involvement in Afghanistan aiming at supporting the government in its National Development Strategy shown in Figure 1.

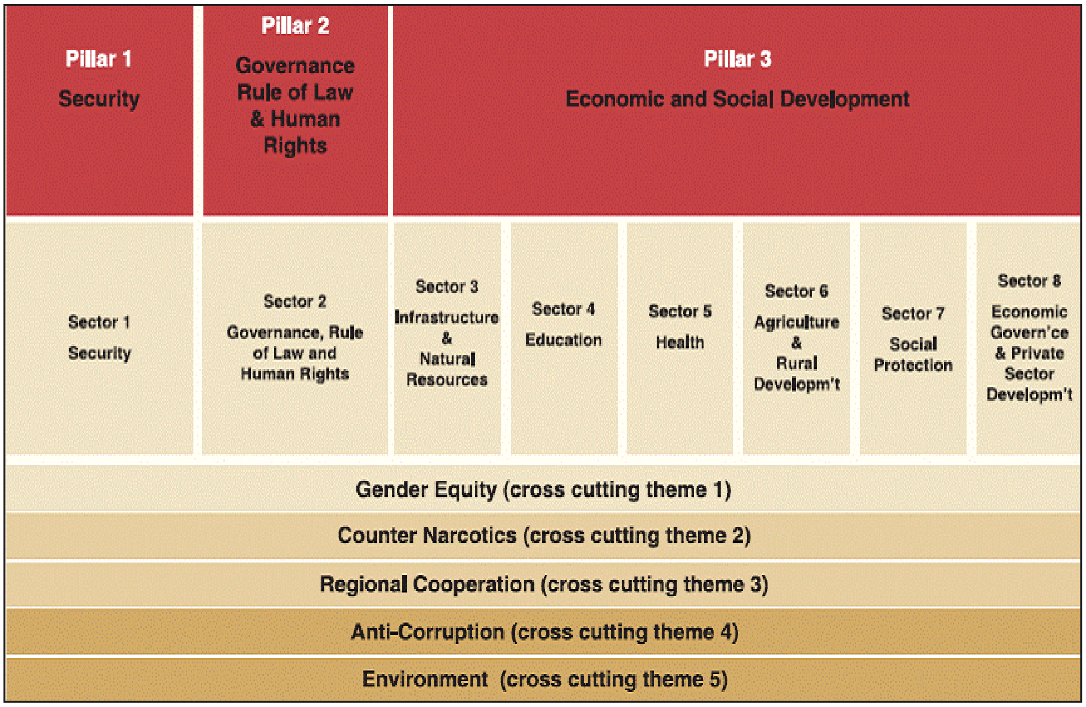


Figure 1: Afghanistan National Development Strategy extracted from (Government of the Islamic Republic of Afghanistan, 2006).

In order to support such a strategy CF has to consider various kinds of interactions with the different actors as presented in Figure 2. Figure 3 summarizes this shift from security mission to these types of mission requiring the collaboration of various experts (military, cultural, political, economical...). From a short-term and specific focus on security, military now needs to consider and not interfere with other government departments and agencies longer-term goals.

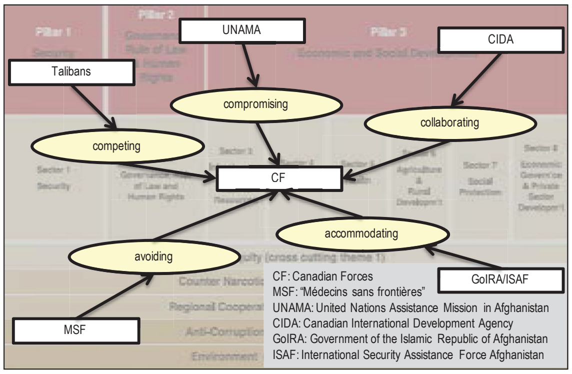


Figure 2: CF interactions in Afghanistan.

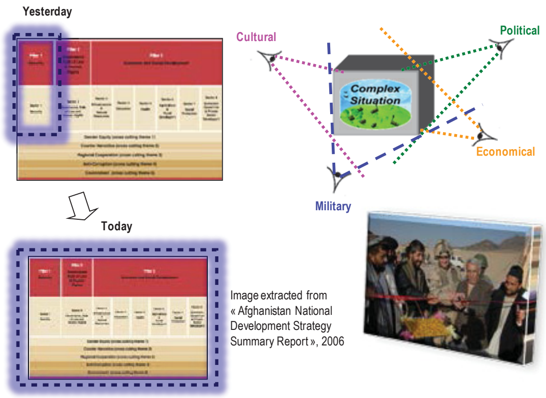


Figure 3: The move from security to complex missions.

2.3 Research goal and objectives

This research aims at studying CSs and imagining a toolset that could help the human to better understand them. Researchers from various research fields and organizations contributed to the effort: software, computer and electrical researchers from DRDC Valcartier and Université Laval (Computer Vision and Systems Laboratory); cognitive psychology researchers from Université Laval (Cognition Distribution Organisation Technologies Laboratory); and operational research scientists from DRDC Operational Research.

From a system engineering point of view, the objective was to conceive and demonstrate the feasibility of a set of software tools, using knowledge representation, simulation and immersive technologies, to assist a team of experts in developing a better understanding of CSs. In terms of cognitive psychology, the objective was to bring evidence that the imagined toolset can augment the understanding of CSs. Finally, operational research brought know-how about military situations and contributed to validate the IMAGE solution and experimental approach, including a realistic definition of a Baseline toolset to be used as the control condition during the cognitive experimentation.

2.4 IMAGE: The proposed concept

The previous pages summarised the problem under study providing the rationales behind the effort reported in the current document while the current section introduces the innovative IMAGE project contribution to address the problem.

Understanding CSs, in order to achieve specific or global objectives, is a very challenging problem that requires various disciplines, new mindsets, methods and tools. The core rationale of the IMAGE project is to provide effective support in this difficult endeavour. Simply stated, the goal is two-fold: (1) increasing understanding of a CS and (2) enabling individuals to share their comprehension. Figure 4 summarises key drivers that guided the elaboration of the IMAGE concept: involving collaboration of various disciplines and applying three principles: (1) Iterative understanding: a common understanding is reached through revision and sharing of successive representations of the situation; (2) Human-driven toolset: the common understanding is above all a human task involving team members using tools their own way; and (3) Synergy of technologies: the common understanding is assisted by the synergy of tools for representation, scenarization, simulation and exploration.

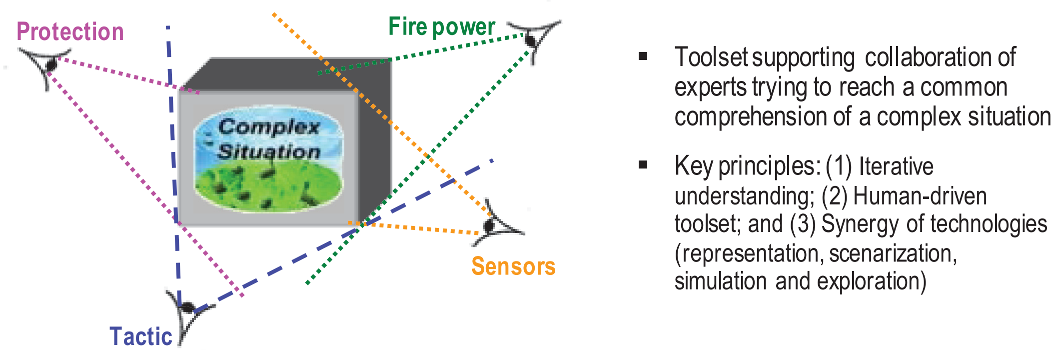


Figure 4: IMAGE.

Figure 5 depicts the toolset concept IMAGE through a wheel. The middle of the wheel illustrates that the Human is driving the understanding effort. The wheel contour identifies the various technologies available while the in-between spiral represents the iterative development of a comprehension from individual perspectives evolving to a common understanding of the CS.

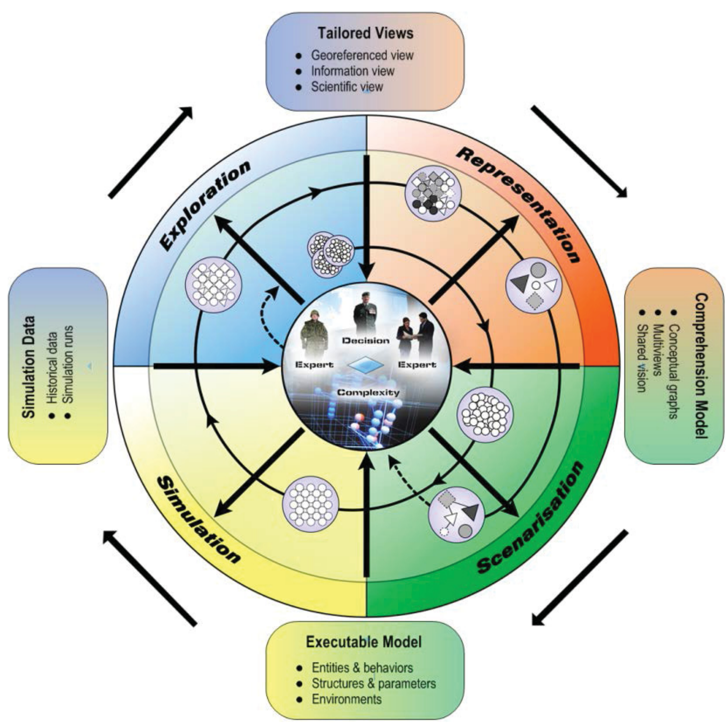


Figure 5: The IMAGE Wheel.

Following, this high-level concept and using their different backgrounds with a will to think beyond a purely reductionist mindset, IMAGE researchers have identified capabilities that each of the modules can bring: Representation (I-REP), Scenarization (I-SCE), Simulation (I-SIM) and Exploration (I-EXP).

I-REP assists experts in recording any piece of information they wish to include, and asserting any type of links between them as well as inferring conclusions from this set of beliefs. Each expert should be able to create an individual comprehension model and share it with other experts. After some collaborative work, the I-REP team efforts result in a common comprehension model i.e. a set of graphs representing the CS.

I-SCE supports agent-based modelling of the CS. Starting from the common I-REP comprehension model, I-SCE assists an expert in working out an executable model. Comprehension model items relevant for a simulation are isolated and standardised using a predefined scenarization vocabulary. Required details, such as properties and behaviour implementations are added to produce an executable model.

I-SIM necessitates a very-high level of interactivity enabling investigation of the space of variables in order to test hypotheses about the CS. I-SIM uses as input the I-SCE executable model and enables an expert to initiate and control parallel simulations by changing parameter values and allowing comparison from a what-if bifurcation points. Such work generates huge sets of data.

I-EXP ultimate goal is to allow an expert to invent views as the comprehension progresses. In particular, it aims at extracting useful knowledge from large datasets generated by I-SIM. It combines visualization and rich user interactions allowing an expert to explore such huge datasets. The resulting views are inputs to enhance comprehension models worked out by I-REP.

Some prototyping has taken place for each module. Tests using evolving versions of both the scenario and the prototype were conducted during the project. The IMAGE Version 1 (V1) prototype, a partial implementation of the concept, was used to conduct a cognitive experimentation to get significant results about how much such automation augments individual's comprehension. IMAGE V1 was a partial implementation of three (3) modules: Representation, Simulation and Exploration. Two key capabilities, collaboration and scenario creation, were voluntarily put aside in this version for feasibility matters of the cognitive experimentation within the project time and budget constraints. It was too costly to create a collaboration experimentation and impossible to give participants enough time to perform some scenarization.

The next version, IMAGE Version 2 (V2) prototype was not but aimed at demonstrating the feasibility of more advanced functionality including collaboration, scenarization and immersion. The I-REP module was enhanced with collaboration capabilities while the I-SCE module was created from I-REP pieces and the Eclipse software development environment. I-SIM was not really modified but was feed with a different scenario and its user interface was integrated in the I-EXP immersive capability. Unlike I-REP, I-EXP V2 is not an enhancement of I-EXP V1 because it was built with a different architecture enabling advanced immersive virtual exploration.

3 The case study: Land convoys facing insurgents

This section presents the two scenarios used to study CSs. They are both about convoys facing various threats. More specifically, the first one (Bernier & Rioux, 2010) was developed to support the development of IMAGE V1 and the experiment aimed at measuring its efficiency while the second one was used to develop IMAGE V2 and test its feasibility. These CSs are described early on in the report in order to support the presentation of IMAGE tools. They comply with the following requirements (Mokhtari et al., 2007):

- Completeness: the CS description is presented at a suitable level, the immediate intent pursued by users of IMAGE is clearly identified and sufficient background material is provided;
- Complexity: the CS must exhibit typical complexity behaviours such as those described in Chapter 5. Because of human experiment constraints, it must be possible "to understand" such a CS within a few hours of IMAGE use. The CS must also be intellectually interesting and challenging for the user;
- Relevance: the CS must be of significance for the Canadian defence context;
- Feasibility: the CS could be easily implemented; and
- Innovation demonstrator: the CS should help to demonstrate the plus-value and the assessment of the proposed tools.

3.1 Background

Canada's role in Afghanistan was characterized by insurgents' attacks using suicide vehicle bombs, improvised explosive devices (IEDs), and rocket propelled grenades (RPGs). Most of these attacks happened during convoy missions. The CS proposed here is inspired by these events, previous publications (Ayvaz, Dere, & Tiah, 2007; DeGregory, 2007) on counter-insurgency and IED, and feedback from military personnel with convoy experience. It should be noted that some elements were adapted to comply with design requirements.

The convoy attack problem can be mitigated by vehicle protection. Although such measure is likely to succeed on the short term, adversary adaptation could reduce its benefit. For instance, a large quantity of armour under the vehicle improves protection from IED but, if it also slows down the vehicles, it could facilitate insurgents' task with RPGs if side armour vehicle is not increased. The latter can also be improved but it would be detrimental to another aspect. Eventually, insurgents would detect and develop ways to exploit newly introduced weaknesses. It is expected that this process would occur as long as one side is not eliminated. This evolutionary arms race is designated two-sided competitive coevolution (McDonald & Upton, 2005). The two adversaries of this CS, the convoy and the insurgents, behave like complex adaptive systems (Holland, 1992).

This idea of convoy-insurgents coevolution was implemented in two parts, a first one at a lower level of complexity named Tactical and another one at a higher level of complexity named standard operating procedure, or SOP as illustrated in Figure 6. Subjects participating to the experimental validation of IMAGE had to study the CS to get an insight on the dynamic of evolution and to understand how it could affect the integrity and efficiency of convoys. Both levels were made deterministic to reduce variability in the experimental data.

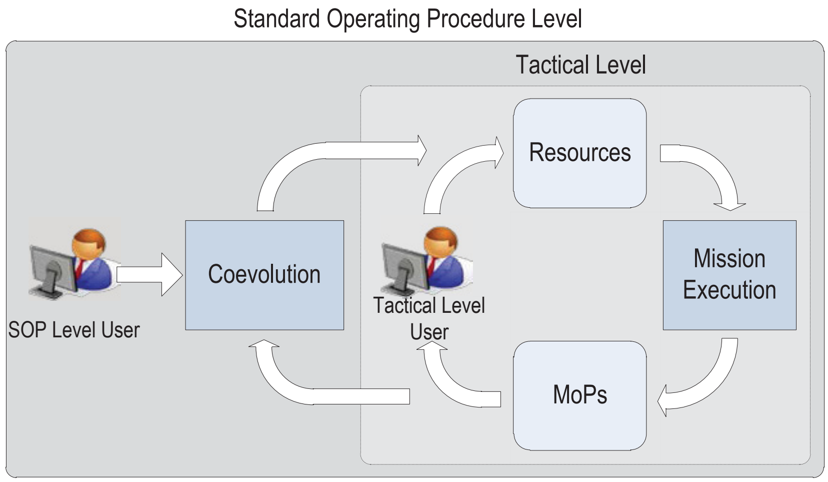


Figure 6: SOP level encompassing the Tactical level for the CS.

3.2 Low level of complexity

The lowest level of complexity of the CS implements the tactical aspect of a convoy mission. A convoy of 10 light armoured vehicles (LAVs) circulates on a road and defends against IED explosions and RPG attacks by insurgents. The CS can be studied according to 11 underside armour thicknesses (UAT) and 11 side armour thicknesses (SAT) for the convoy and 6 IED / RPG allocations by the insurgents. The resulting parameter space comprises 11×11×6 = 726 combinations of variables. Resources must be allocated prior to the execution of the mission, as illustrated in Figure 7. When a mission ends, a set of measures of performance (MoPs) for each side is calculated. During the experiment, participants had to understand the impact of resource allocation on the outcome (i.e. MoPs) of the mission.

Figure 7: Simulation process at the Tactical level.

The Tactical level was designed to be minimally but still complex. The reader must not be misled by the term lowest level of complexity because it still represents a challenging situation to understand as the experimental results will confirm. Its elements of complexity are:

- Counterintuitive behaviour: higher SAT increases convoy integrity since it protects against RPGs but, with too much of it, it reduced vehicle speed such that it makes it more vulnerable to RPGs.
- Rock-paper-scissor dynamic (no one-size-fits-all solution). For instance, an armour thickness that performs well against a given number of IEDs / RPG shooters will perform poorly against a radically different resource allocation by the insurgents.
- Mutually exclusive goals, hence a multiobjective optimization (Jin & Sendhoff, 2003) problem.

3.3 High level of complexity

The high level of complexity of the CS implements the SOP aspects of convoy missions over the long term. The SOP level comprises 100 successive convoys (called iterations). Between each convoy, two optimization algorithms, one for each side, modify previously defined resource allocation values (UAT, SAT and NIED / NRP). The algorithms select the values that maximize the chosen MoP, either integrity or cargo, at the next mission. The result is a coevolution similar to what can be observed in long conflicts. The convoy evolution is customizable in two ways. First, a portion of the transported cargo can be dedicated to reconstruction operations. This portion of cargo improves population allegiance which facilitates intelligence gathering which, in turn, influences positively the evolution of the convoy side. Second, for each iteration, either integrity or cargo must be selected as the variable to optimize. During the experiments, participants could control a few variables: initial resources allocations, priority of the evolutionary algorithm at each iteration, and the percentage of cargo given to the population at each iteration. Participants had to find the combinations of those variables corresponding to the highest possible cumulative values of integrity, cargo and timeliness over 100 iterations. Figure 8 shows the evolution process of the SOP level.

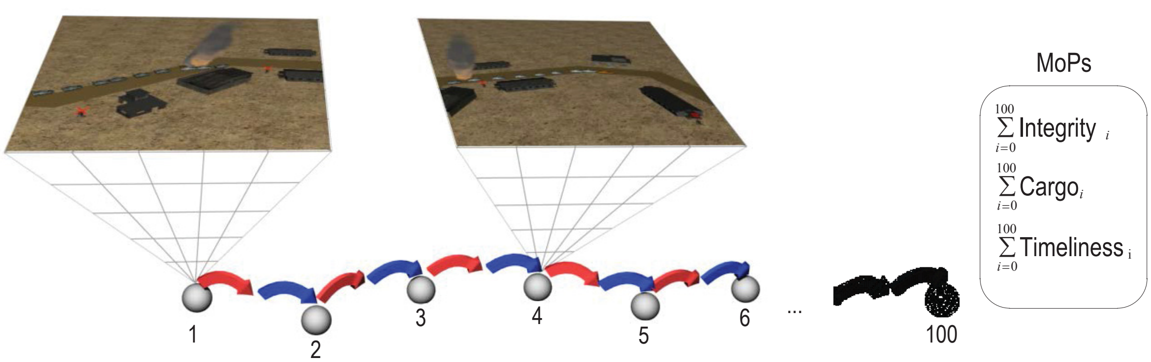


Figure 8: Coevolution process showing the embedded Tactical level (red arrows indicate insurgent evolution while blue arrows designate convoy evolution).

The SOP level is more complex and challenging. Its elements of complexity are:

- Vast parameter space composed of 10^{13} elements.
- Coevolution causing an arm race for which no absolute best solution can be found since the Tactical level is designed to be self-balanced by using an intransitive approach (like the rock-paper-scissor game). Intransitivity and bounded resource allocation lead to cycling in the coevolution (De Jong, 2004).
- Butterfly effect and feedback loop: small changes of SOP leading to dramatically different courses of evolution that would undermine the any plan to optimize the MoPs.
- Non-respect of the Ashby's law of requisite variety (Ashby, 1956) which states that a perfect control of a system is only possible if the number of states in the control system is equal or higher than the total number of states of the system being controlled.
- Counterintuitive behaviour: the integrity is often higher over the long term if optimisation priority is set to cargo instead of integrity.

3.4 CS for IMAGE V2

A second CS was designed for the immersive version of IMAGE (IMAGE V2). The original convoy theme of convoy and counter-insurgency was kept but augmented with a geospatial component. The first addition is a network of roads on which the convoy circulates to transport material aiming at constructing a school. The second addition is a geospatial version of the population allegiance that is negatively affected by close IED explosions and positively affected by convoy circulating in proximity. Convoy commander chooses routes and vehicle types while insurgents choose IED strength and location. An artificial intelligence algorithm plays the role of the commander and the insurgents so the mission can be executed automatically. In addition, the algorithm can be customized or even completely replaced by a user. Finally, every parameter of the CS can be modified. Figure 9 shows the three simulation steps with the actions to be executed. The mission stops when the school is completed, i.e. between 5 to 10 convoy trips.

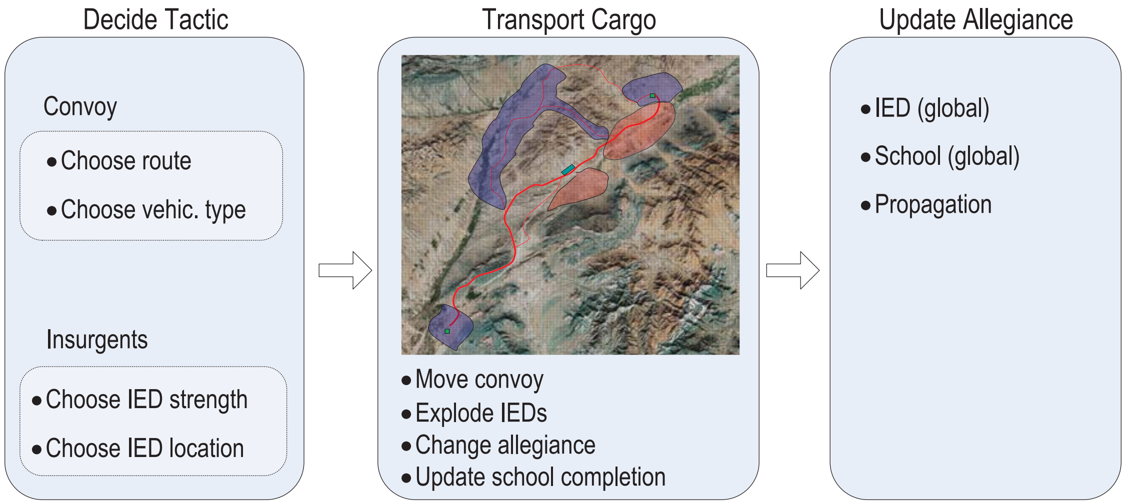


Figure 9: CS of IMAGE V2 showing actions to be executed at each step.

This page intentionally left blank.

4 Representation: Crafting a comprehension model

Understanding appears in the mind of people. We gradually form a mental model of reality, more precisely our perceived reality, which could be different from one person to another. This mental act does not involve any automation. However, to better assimilate a problem or share it with others, we often have the reflex of drawing pictures on a piece of paper, a white board or even a napkin. In the same line of thoughts, anyone can probably remember explaining a problem to a colleague to obtain assistance, but understanding and solving it even before the colleague has time to say a word. Based on this strength of expressiveness, the IMAGE Representation (I-REP) concept consists in expliciting a mental model into a comprehension model.

4.1 The Representation concept

As illustrated in Figure 10, the I-REP concept aims at supporting the production of individual comprehension models and their sharing with other team members. This concept is based on two hypotheses. On one hand, such a task will increase comprehension of the user. The process of expressing or formalizing a mental model should have a positive feedback on the user's own understanding. On the other hand, letting other team members look at an individual's comprehension model is the best way to help them to understand this individual view (or mental model) on the situation and consequently should lead to a better common team comprehension.

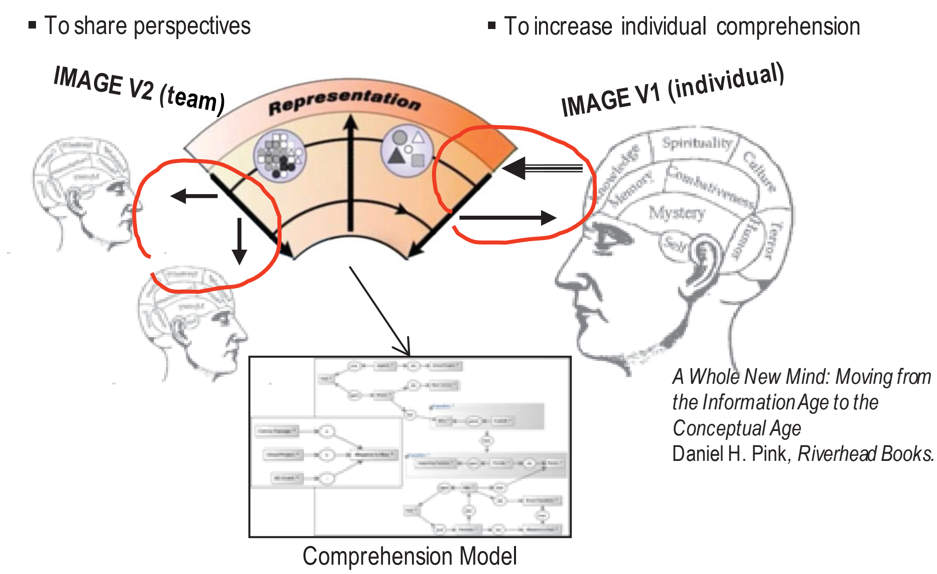


Figure 10: The I-REP concept.

As in solving a puzzle, understanding a CS does not occur in a predetermined sequence. People link pieces into sets, eventually link some sets to finally get the big picture. It is similar to a police inquiry or a military intelligence effort trying to discover enemy intents. Considering that people are not good at calculations and can hardly handle more than seven concepts at a time, an I-REP toolset should assist users to record any piece of information, establish any type of link between these pieces, and infer conclusions from this set of beliefs. Comparing and merging information are other examples of advance operations users would need.

The I-REP concept has been studied through two prototypes: I-REP V1 and I-REP V2. I-REP V1 was part of the IMAGE V1 cognitive experiment focussing on individual understanding while I-REP V2 was part of the IMAGE V2 feasibility effort and focussing on the collaboration part. The I-REP prototypes (also called IMAGE-CoGUI) were developed from the CoGUI tool. The latter software is founded on the conceptual graph research efforts of Chein and Mugnier (2004) as well as on work by Genest (2012).

4.2 Related work

As mentioned above, the I-REP concept aims at supporting specialists in the production of individual comprehension models and their sharing with other team members. Let's look at related works from three angles: knowledge representation, collaboration tools and software usability.

Knowledge representation

The I-REP concept introduced above had to be implemented with a graphical notation enabling sharing of concepts between team members. Graphical notations are used by various disciplines but often for a single specialty: music, architecture, electrical engineering, mechanical engineering, software development… The ones of interest for the current research had to be shared by experts with different backgrounds, a bit like a common speaking language: it allows talking about anything but need to be known to all experts. Therefore such a graphical notation has to be a generic formalism facilitating the representation of any CS concepts but more particularly interactions, which are keys to understand the dynamics. Visual knowledge representation notations were of prime interest closely followed by system dynamics formalisms. Conceptual graphs were selected after looking at various alternatives. Approaches such as Conceptual mapping (O'Connor, Johnson, & Khalil, 2004), Causal Loop diagrams (Kirkwood, 1998) and Influence diagrams (Curtis & Dortmans, 2004) are generic and had been mainly used to explicit or to measure understanding of complex systems. However, they miss the richness of notations that were created to handle and present knowledge such as behavioural graphs (Bazan, Peters, & Skowron, 2005), adaptive fuzzy cognitive maps (Dickerson & Kosko, 1994) and conceptual graphs (Sowa, 1998). The mixed of cognitive maps and conceptual graphs proposed by (Aissaoui, Genest, & Loiseau, 2003) had particularly gotten our attention for this study. This approach combines cognitive maps, enabling visualization and computation for influence on a goal, to conceptual graphs used to precisely define each notion used in the cognitive map. Although very interesting, we did not invest effort on propagation of effects and therefore were not requiring a real concept map computation capability. Instead, we used directly the conceptual graph notation to represent influences.

Collaboration

The notation had to be generic but at the same time it needed to support the various independent specialist perspectives in the spirit of an IEEE standard for architectural description of software-intensive systems involving various stakeholders. The IEEE Std 1471 standard (Hilliard, 2000) is intended to stakeholders who need to understand the architecture of a system from different perspectives: owners, end users, developers. This approach was of interest not for its notation but for its way of dealing with multiple perspectives over a same reality, the architecture of a system. By analogy, the *architecture* corresponds to the CS to be understood and an *architectural description* (a document describing the architecture) corresponds to the common understanding worked out by the specialists. The *architectural description,* the common understanding, is intended to stakeholders and split in a set of views. A view is presented from a specific viewpoint covering a set of concerns of interest to a category of stakeholders.

Various studies were conducted around the collaboration topic among which many investigated learning skill and measurements: (Milrad, 2002), (Mohammed & Dumville, 2001), (Lewis & Sycara, 1993), (O'Connor & Johnson, 2006). These researches have provided a lot of good inputs to help defining the kind of features needed by a team. We did not found research on interactive collaboration centered on conceptual graphs. Many researchers have look at operations over conceptual graphs, such as Chein and Mugnier (2004), Corbett (2003), and Genest (2012), while others, such as Ribière, Matta, and Cointe (1998) and Dieng and Hug (1998) have investigated the notion of viewpoint. They could be eventually exploited for collaborative work context but none directly addresses interactive collaboration.

Usability

As emphasized by the second principle underlying the concept IMAGE, the development of a common understanding is above all a human task. A technology can be excellent but is useless if users cannot exploit it properly. Such situations are often the main reason explaining failures of information systems implementation. Therefore, the usability subject is at least as important as the knowledge representation and collaboration ones.

Although studying software program comprehension and not CS understanding, Storey (2006) identified tool requirements applicable to our effort such as searching and multiple views in addition to cognitive theories such as top-down, bottom-up and opportunistic comprehension. Another interesting observation specifically about conceptual graph is coming from Polovina (1993) who is stating that given that conceptual graphs theory is the most technically yet visually advanced knowledge-based formalism, visualisation is the challenge for complex problems such as strategic ones. Other relevant works are providing user-interface design guidelines such as Wickens, Gordon, and Liu (2004) 13 principles.

In summary, testing the I-REP module and the overall IMAGE concepts is a very challenging task since the implementation is critical. Good knowledge representation and collaboration functionality has to be provided but moreover it has to be usable. Therefore, IMAGE experimentation was not only performed to test the concept but also to improve its implementation.

4.3 Graphical notation: Representing a situation

Figure 11 provides an example of the graphical notation of the conceptual graph formalism (Sowa, 1984) variation used in I-REP V2[1]. Rectangles are concepts, while ellipses are conceptual relations. A concept has two parts, a type label before the colon and a marker after the colon. The type label represents the type of entity the concept refers to, while the marker refers either to the generic marker "*" or identifies actual individuals, also called referents. Arcs pointing toward or away from an ellipse mark arguments of the relation. The arrow orientation is devoid of semantics and serves only to ease graph reading[2]. A dotted line linking two concepts is called a co-reference and means that they actually represent the same thing. A co-reference is not valid if both concepts are using different referents. A concept can also be detailed using a nested graph, as shown in Figure 11, where the concept "IED Event:*" is detailed with the graph "IED:*→ patient→Explode:*". This part of the graph means that an IED Event is actually an IED that explodes. The IED concept (a kind of entity) is in relation with the Explode concept (a kind of behaviour) through the relation patient that links a passive entity, a patient, to a reaction behaviour. The notation can also represent dynamics of a situation such as the positive influence relations in "School Project:* → + →Allegiance to Blue:*".

[1] As the project was progressing, needs for lightening conceptual graphs were identified. As a consequence, the graph notation was extended to reduce the quantity of shapes and lines (Lizotte & Rioux, 2010).

[2] In the original conceptual graph notation, an arc pointing towards an ellipse marks the first argument of the relation, while one pointing away from an ellipse marks the last argument. If a relation has only one argument, the arrowhead is omitted. If a relation has more than two arguments, the arrowheads are replaced by integers.

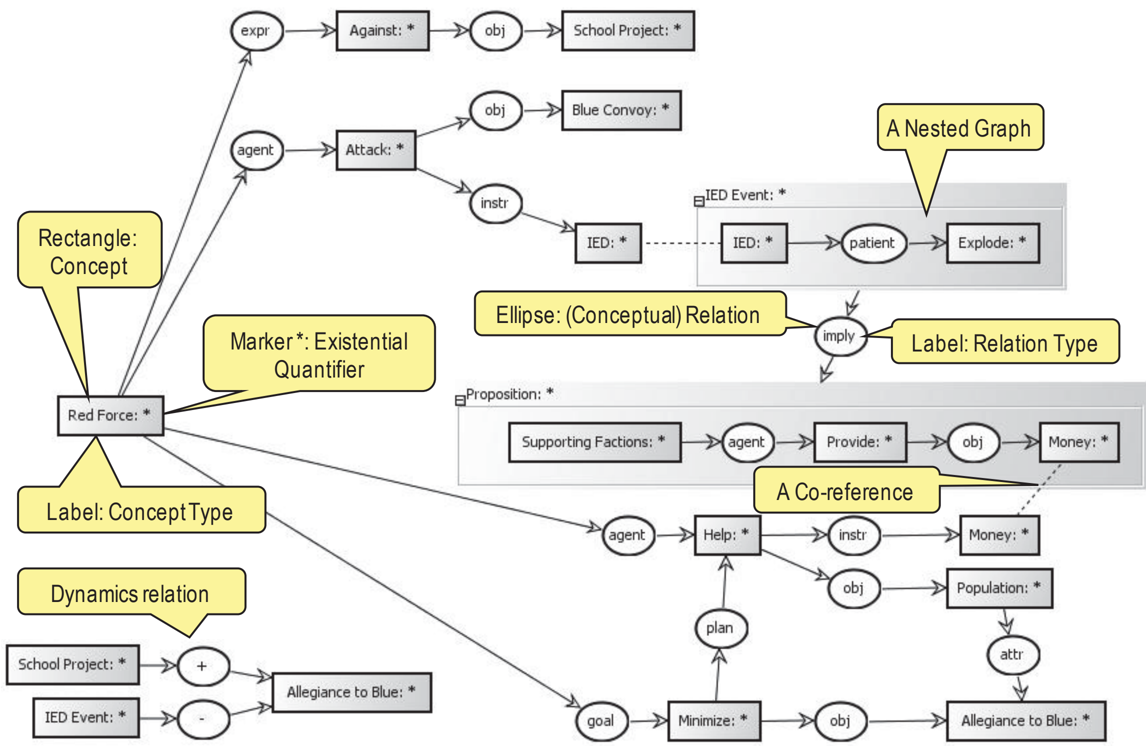


Figure 11: Conceptual graph formalism.

4.4 Vocabulary: Formalising a terminology

In the context of IMAGE, which essentially uses the definitions of Chein and Mugnier (2009), a vocabulary is a simple ontology composed of a concept type set and a relation type set, also called hierarchies. Both sets are partially ordered by “is-a” (shown as an arrow) relations. In addition, a relation type has a signature specifying the arity of the relation and the “maximal” (hierarchically highest permitted) concept type of each argument. Using both, a textual tree and a graph layout, Figure 12 presents an example of a vocabulary implemented with I-REP V2.

The comprehension vocabulary explicitly structures and labels concept types and relation terms used in the graphs. The creation of such a vocabulary is mandatory if one intends to use the automated reasoning power of conceptual graphs e.g. validating, querying, inferring, joining...

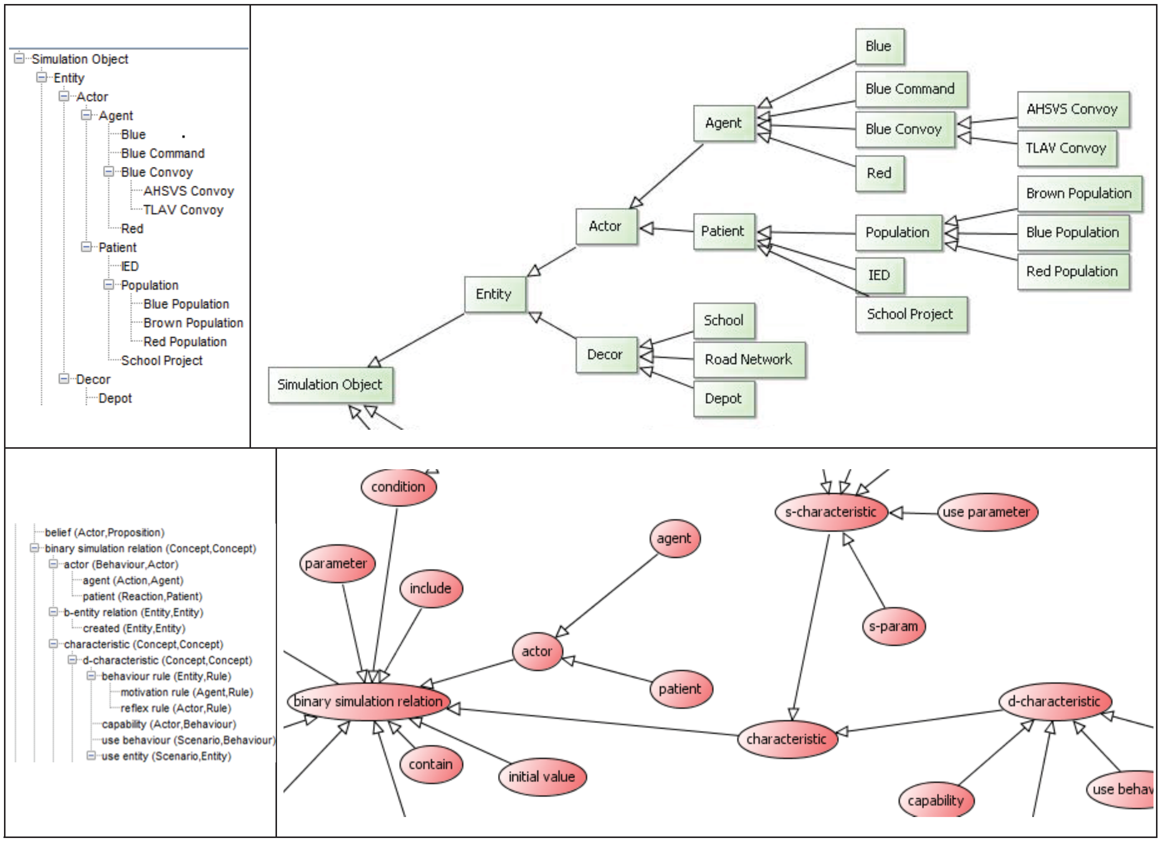


Figure 12: Vocabulary.

4.5 Comprehension models: The I-REP result

The main output of I-REP is one or more comprehension models where a model is set of graphs and a vocabulary (concept types and relation types) as illustrated in Figure 13.

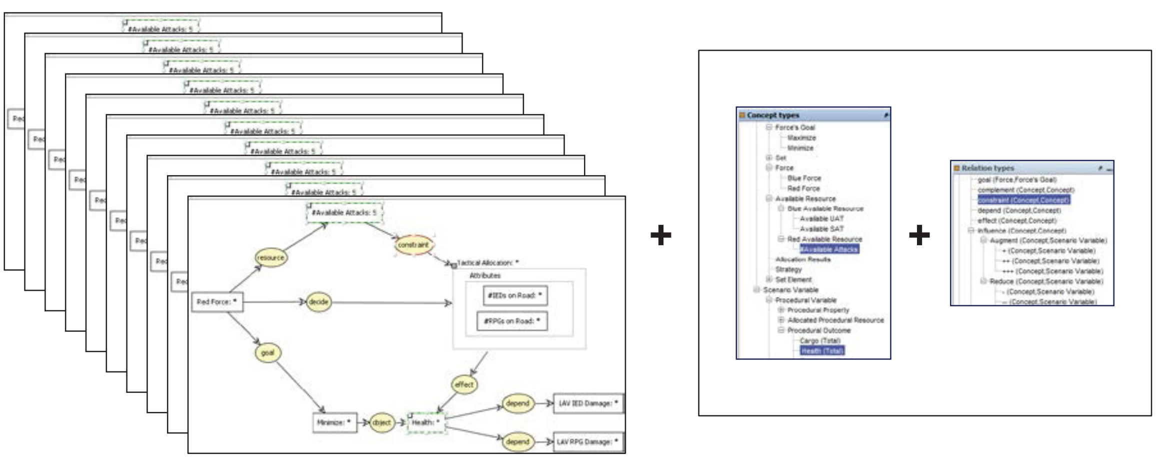


Figure 13: A representation model.

4.6 Unified views: Merging individual perspectives

As mentioned previously, a sound approach to develop a common understanding is to let each team member express an individual comprehension that all others can appreciate and, to enable collaboration around the various resulting perspectives. In an ideal process, the team would iteratively work out a single common model of the CS.

But properly grasping what another individual means about a CS implies some challenges. We have identified three. The first one is the capability for an individual to explicit his / her thought. Although it is a key to success, it was not directly addressed in the current study, but experimental results provide some hints. This gap in the current work could now be address in future studies starting from these preliminary results.

The second challenge is terminology. People often use different words to designate the same concept but worst, they frequently use the same word (or an acronym) to designate different concepts. For instance, do “insurgents”, “enemy” and “red force” really mean the exact same thing or do they represent different flavours of the same threat? Does mission mean a group of people (i.e. a diplomatic mission), a military operation or a whole-of-government set of objectives? Does HA mean humanitarian assistance or a holding area in the military sense? I-REP V2 addresses this challenge through tools aiding the development of a common vocabulary.

Merging individual perspectives is the third challenge. Ideally, at the end of the collaborative effort, the resulting common model of a CS result should not include any redundancy or contradiction. On the other hand, pieces created by various team members should complement each other. The I-REP V2 prototype toolset tackles this problem with functionality to query and unify a set of graphs created by different team members. It also enables them to find similarities, discrepancies and incoherencies between conceptual graphs.

4.6.1 Developing a common vocabulary

I-REP V2 uses two categories of vocabulary items: private and public. Each team members owns a private vocabulary that others cannot access. A vocabulary owner can create, modify, delete and link private items. However, all team members share a common vocabulary developed in collaboration using a lock-unlock approach. In addition to the read-write access to a private vocabulary, all team members have a read-only access to the common vocabulary. A team member can temporally lock the common vocabulary to create, modify, delete and link public items. All others keep their read-only access but cannot perform any modifications.

Figure 14 shows a concept type vocabulary including public and private items. The left hand side of the figure presents the vocabulary using a textual tree view where public items are in bold characters and private items are in regular characters. The right hand side of the figure shows the vocabulary using a graph layout where public items have a white background and a bold contour and private items have are coloured background and a regular contour.

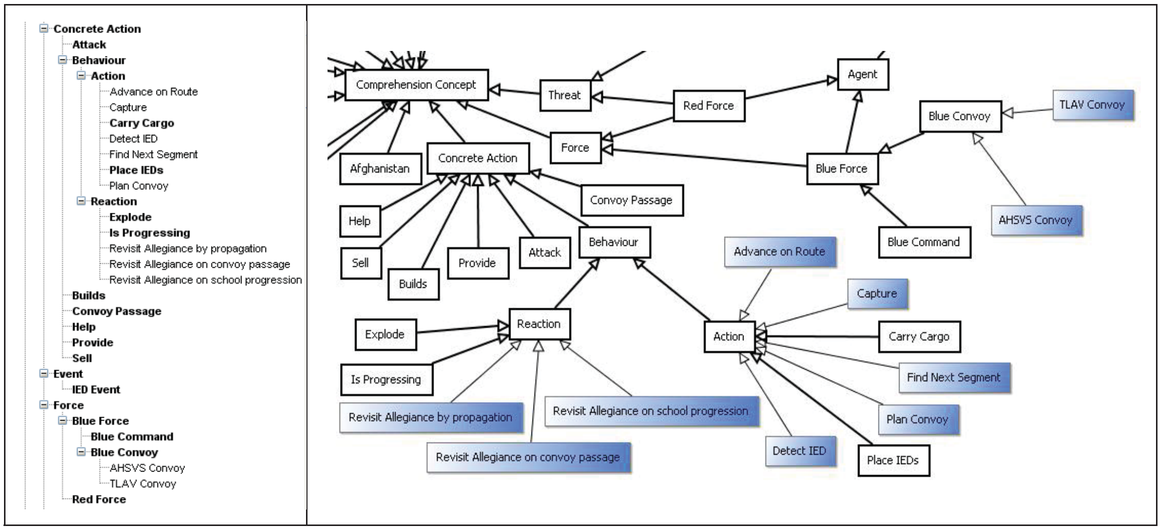


Figure 14: Common vocabulary.

As shown in Figure 15, concepts and relations in an I-REP conceptual graph fit in three categories of graph elements: draft, private and public. Concepts or relations referring to public vocabulary items have a bold contour while the graph elements referring to private vocabulary items have a regular contour. Graph elements that are not referring to vocabulary items have a dash contour.

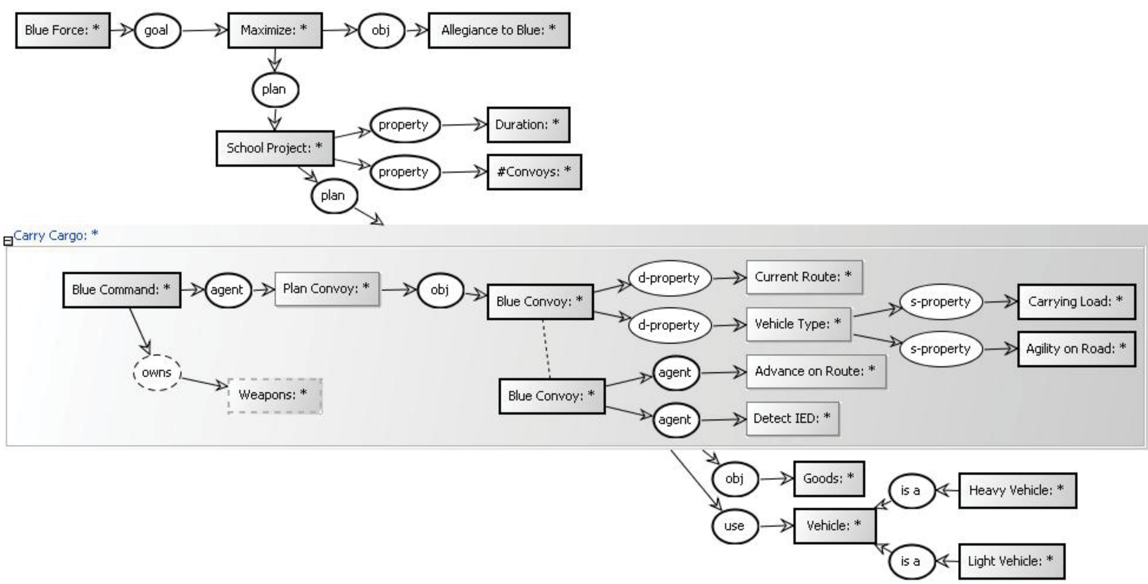


Figure 15: Draft, private and public graph elements.

4.6.2 Querying a set of conceptual graphs

Figure 16 depicts the usage of the conceptual graph power to find patterns in a knowledge base. The belief graph piece (a) asserts negative influences "–" between four variables. The rule graph (b) asserts that a positive influence "+" between two concepts A and B can be deduced from two negative influences “–” between A and X and, between X and B. The vocabulary extract (c) asserts, among other things, that the influence “+” is a kind of influence “augment”. Finally the query graph (d) is looking for the scenario variables that augment the cargo. Figure 17 presents a generated result graph highlighting the two results.

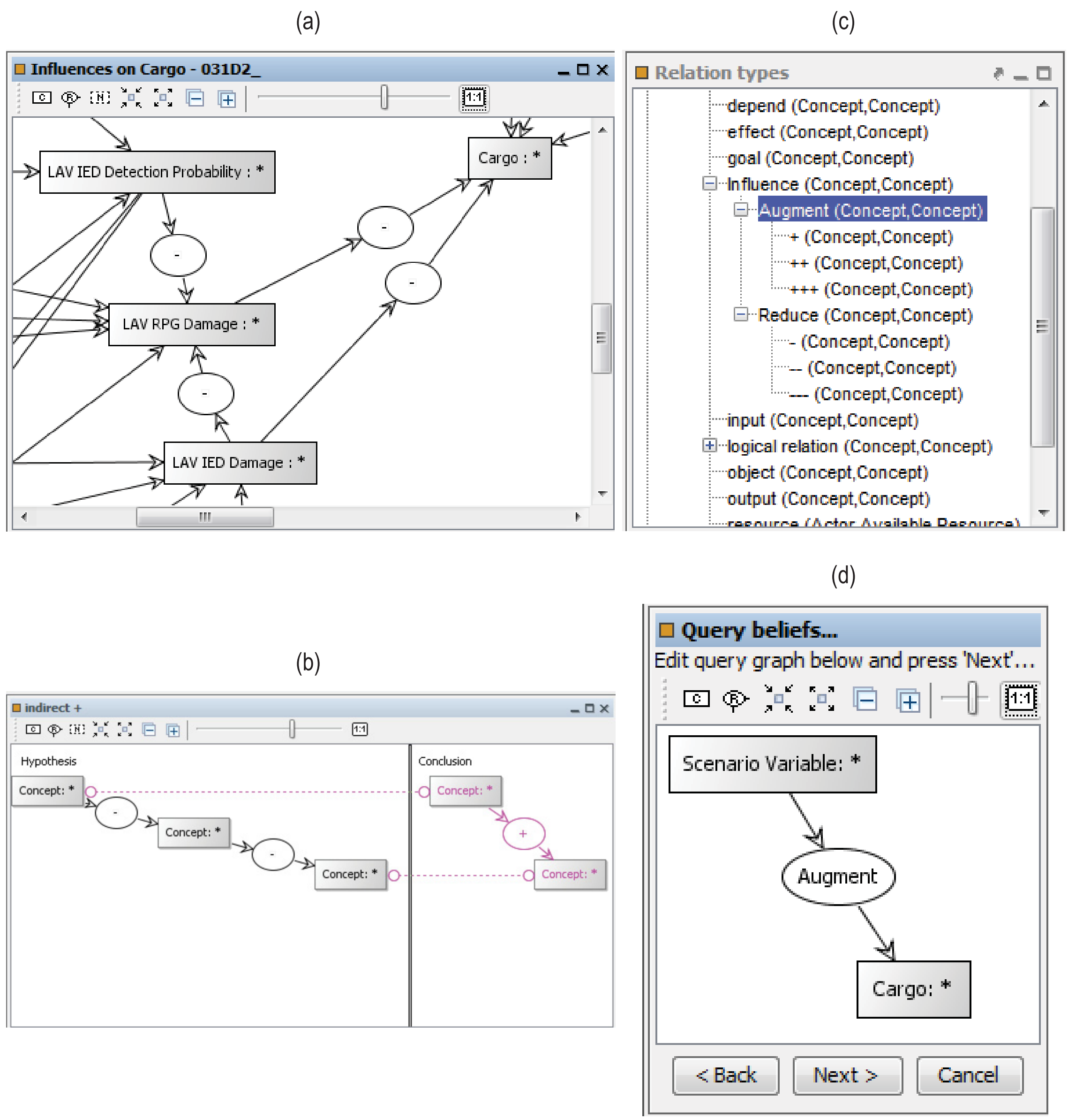


Figure 16: Belief query.

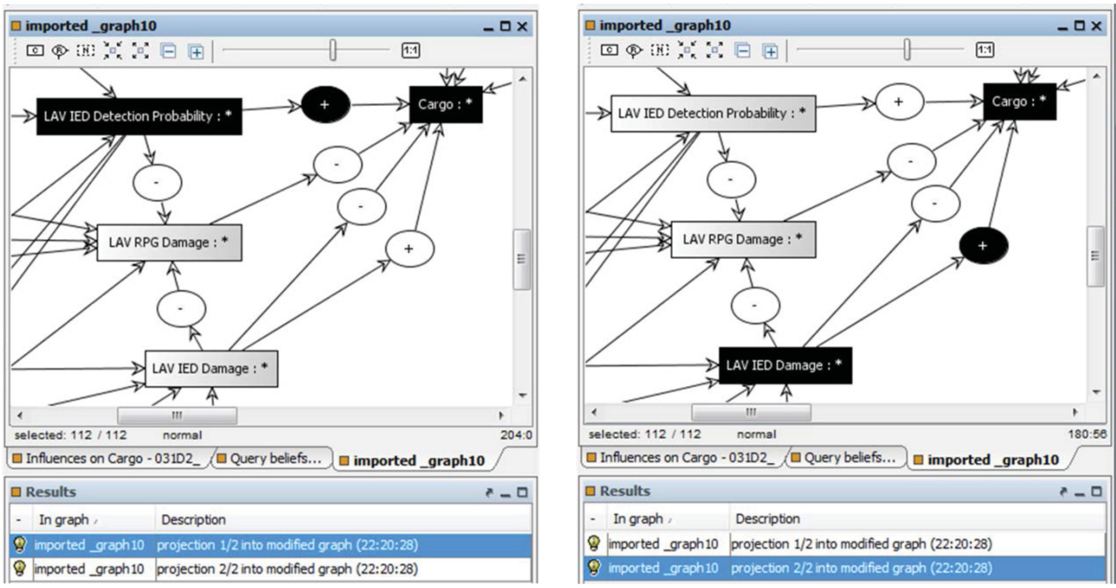


Figure 17: Query results.

4.6.3 Unifying a set of graphs

The first common step in a collaborative understanding environment is to let each expert work out an individual understanding. Each expert identifies the main concepts and relations involved in the situation under study from a specific domain viewpoint. The example used here involves three experts: Blue, White, and Red. Each one has a personal perspective on the situation that can be composed of many graphs but limited to three for the demonstration: Figure 18, Figure 19 and Figure 20.

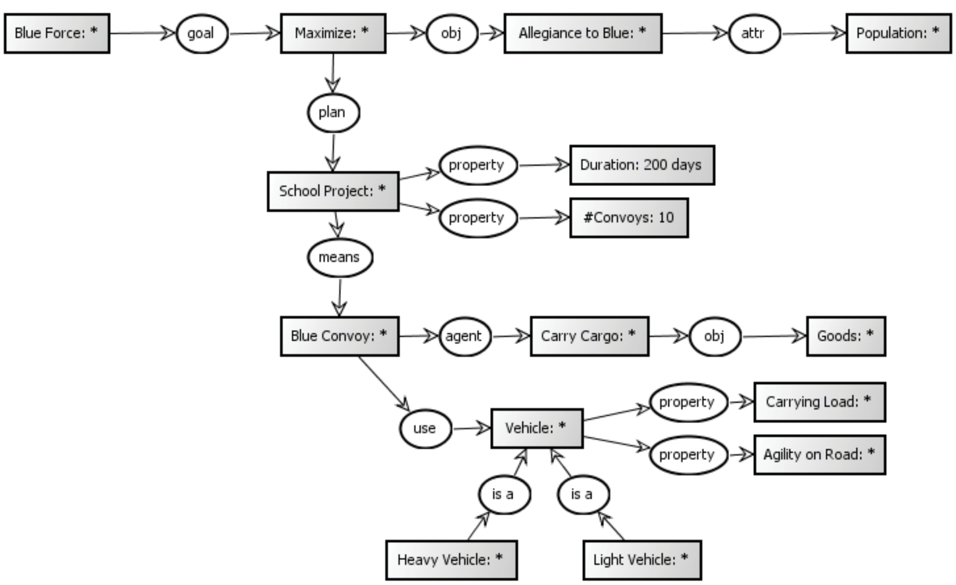


Figure 18: Blue perspective.

For some of their concepts, these graphs are referring to the same public concept types. The Blue and White perspectives share “Allegiance to Blue” and “School Project concept types”.

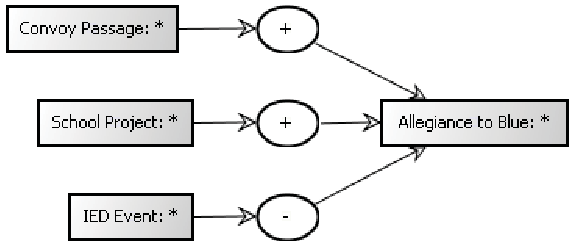


Figure 19: White perspective.

The White and the Red perspective share these two items and the “IED Event” concept type.

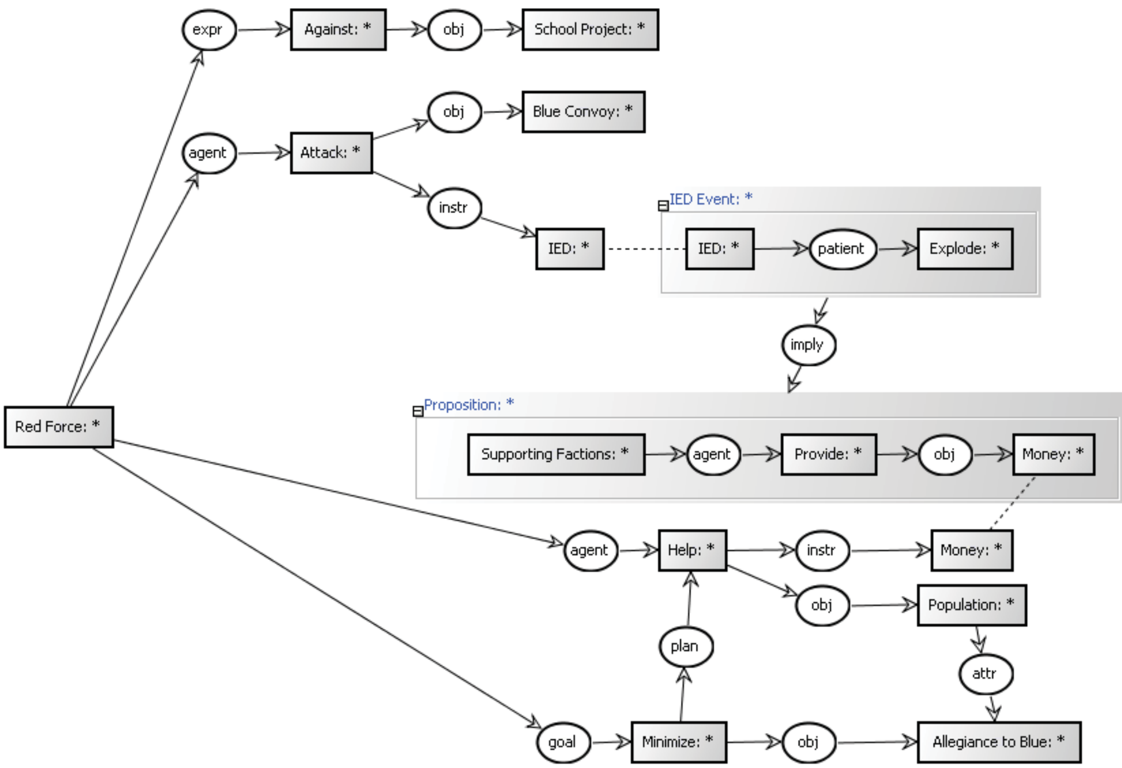


Figure 20: Red perspective.

Two graph concepts referring to the same concept type can designate the same object. Obviously they certainly do if they use the same referent but it is also possible if one or both of these concepts is (are) using the generic marker “*” (representing the existential quantifier). However, two graph concepts referring to the same concept type do not necessarily design the same object. Notwithstanding this fact, it can be very useful to hypothesize possible unification of matching concepts to study a situation. The I-REP view feature does that: A view presents a hypothetical graph resulting from the automatic unification of a set of belief graphs selected by the expert. Then, it is up to the expert to create or not a belief graph from this hypothesis, where all matching concepts are unified. An I-REP view can also use filters on the unified graph to find subgraph patterns. Figure 21 presents a set of three filters applied on the unified graph of Figure 22, which illustrates a view automatically built from the previous Blue, White and Red perspective graphs.

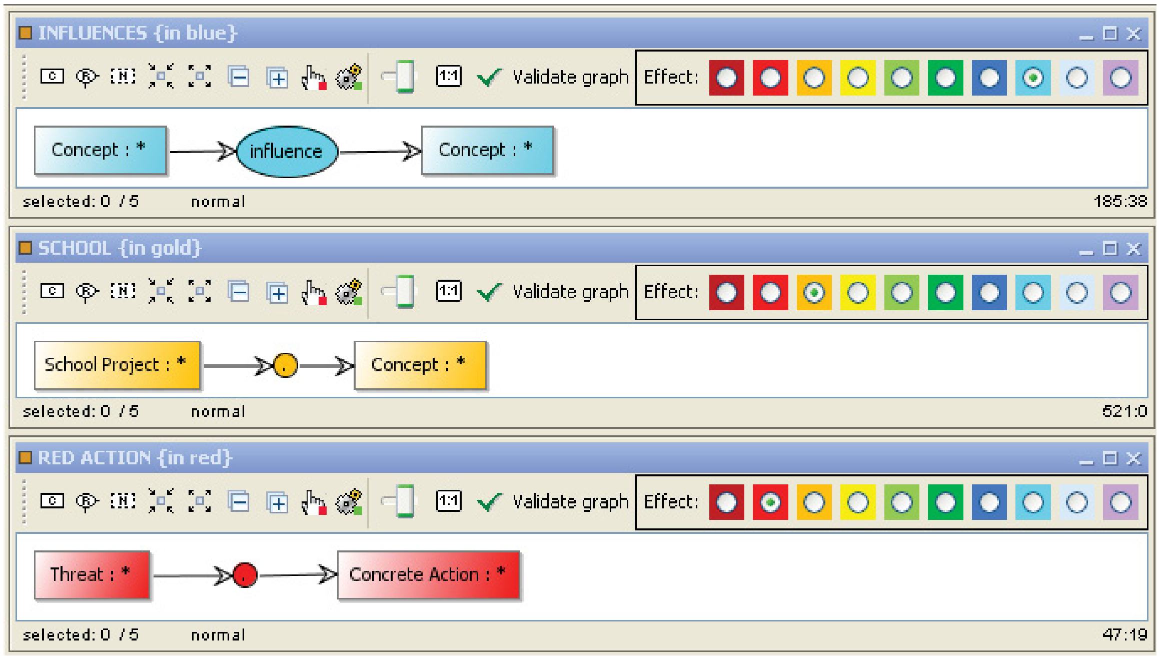


Figure 21: Filter graphs.

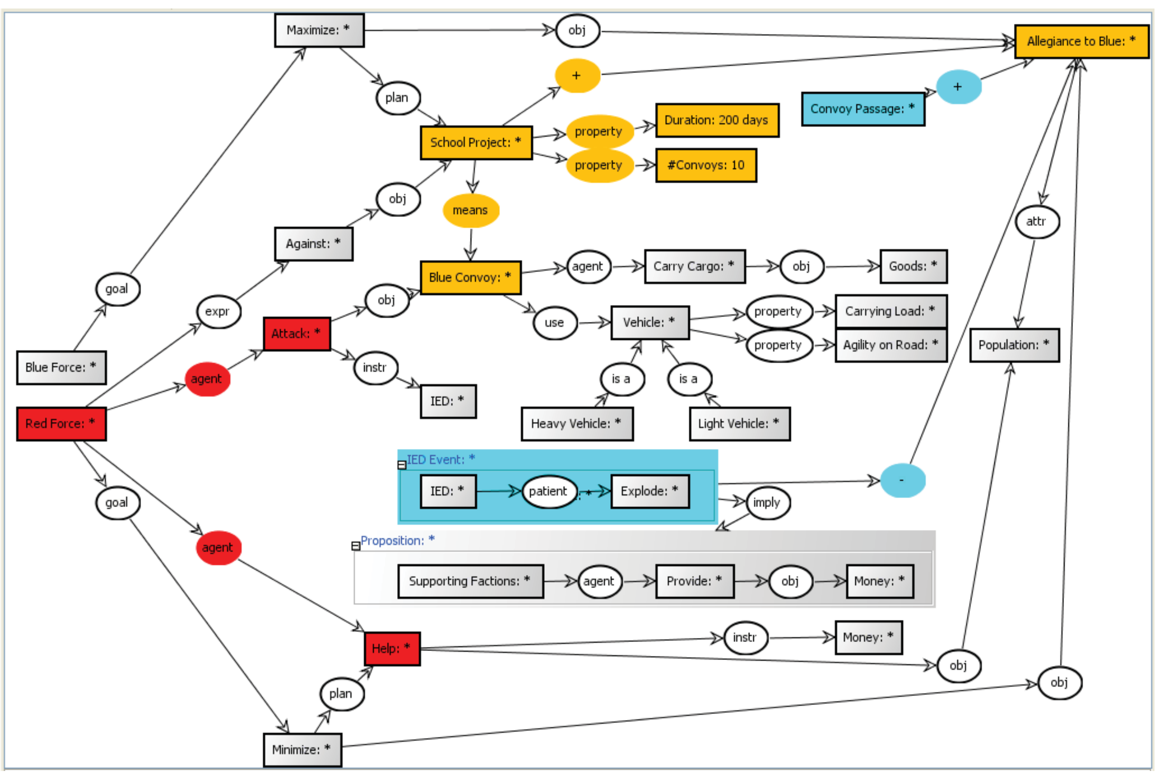


Figure 22: Filtered unified graph.

4.7 Software architecture: Reusing an existing tool

I-REP V1 and V2 prototypes were built from an existing tool called CoGUI developed in the context research efforts of Chein and Mugnier (2004) efforts as well as work by Genest (2012). The CoGUI tool was selected following a search in the conceptual graph literature and over the internet. Actually CoGUI is a Java Graphical User Interface wrapping the C++ CoGITaNT - Conceptual Graphs Integrated Tools allowing Nested Typed graphs – class library (Genest, 2011) that is actually the engine really implementing the conceptual graph theory and manipulating the conceptual graphs. The mandatory selection criterion of the I-REP tool evaluation process was the access to the source code enabling its adaptation to the specific project needs. Seven conceptual graph tools passed this mandatory criterion. Then, other various criteria were considered: available features and usability, existence of a development community, documentation, storage support, support of collaboration …

The main strengths of CoGITaNT are its support of all conceptual graph theory operations as well as extensions such as inference rules and typed nested graphs. On the user interface side, CoGUI is offering interesting features to create and manage and ontology and use it to create, validate and modify conceptual graphs.

The main I-REP V1 development effort was focussed on achieving a high level of robustness and reliability to face the cognitive experimentation. A link with the I-SIM module was also developed to automatically generate graphs presenting inputs and outputs of a simulation. For I-REP V2, a very important effort was required to implement: (1) a client-server enabling collaboration for the vocabulary development and graph sharing and (2) views supporting automatic graph unification and filtering of the resulting unified graph.

This page intentionally left blank.

5 Scenarization: Generating an executable model

A simulation based on current beliefs about a situation is either a mental act, such as a visualisation technique, or an actual exercise human performed to better approach the future. To increase their grip on a situation, the military use war-gaming to fake dynamics between friendly and adversary forces. In the same line of thoughts, the IMAGE concept includes pieces to address such an approach aiding to understand a situation. The preparation of a simulation is what IMAGE Scenarization (I-SCE) is about. The I-SCE concept consists in transforming the I-REP comprehension model into an executable model.

5.1 The Scenarization concept

Figure 23 illustrates the I-SCE concept that aims at generating an executable model of a situation. The user selects concepts and relations from the comprehension model that are relevant for a simulation. Each selected elements is standardised using a predefined scenarization vocabulary while useless items are put aside. Details required for a simulation, such as properties and behaviour implementations are then added to produce an executable model.

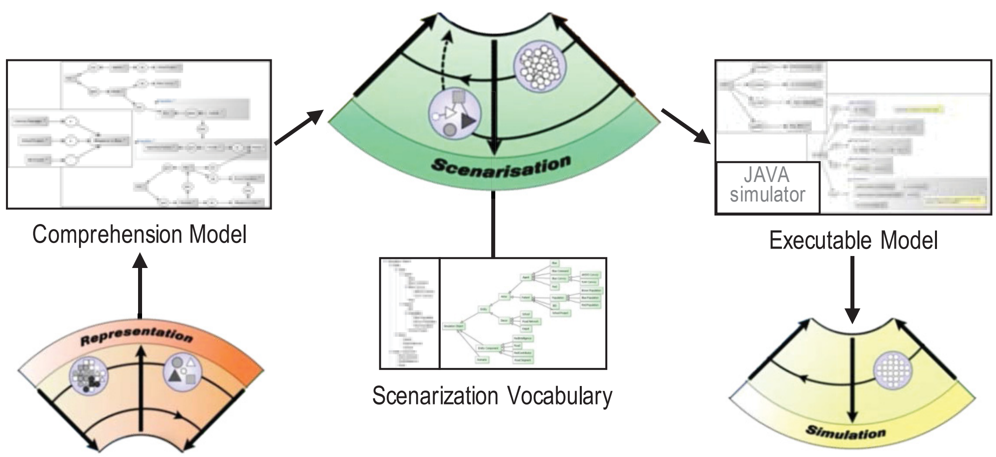


Figure 23: The I-SCE concept.

A normal usage of the I-SCE concept is depicted in Figure 24. The overall approach enables the incremental and iterative representation of a problem and its translation into an executable model. The user is handling three models during this spiral process. The comprehension model is used to express the problem without considering the simulation needs: the user is free to express any knowledge of interest. The scenarization model is an intermediate model including simulation objects (see Section 5.3) and their specifications (see Section 5.4) that are required to generate the source code of fundamental simulator software package. Simulation dynamics implementation details are inserted (see Section 5.5) in the generated source code to produce the executable model which execution results become inputs to previous activities and may lead to retroactions on the three models.

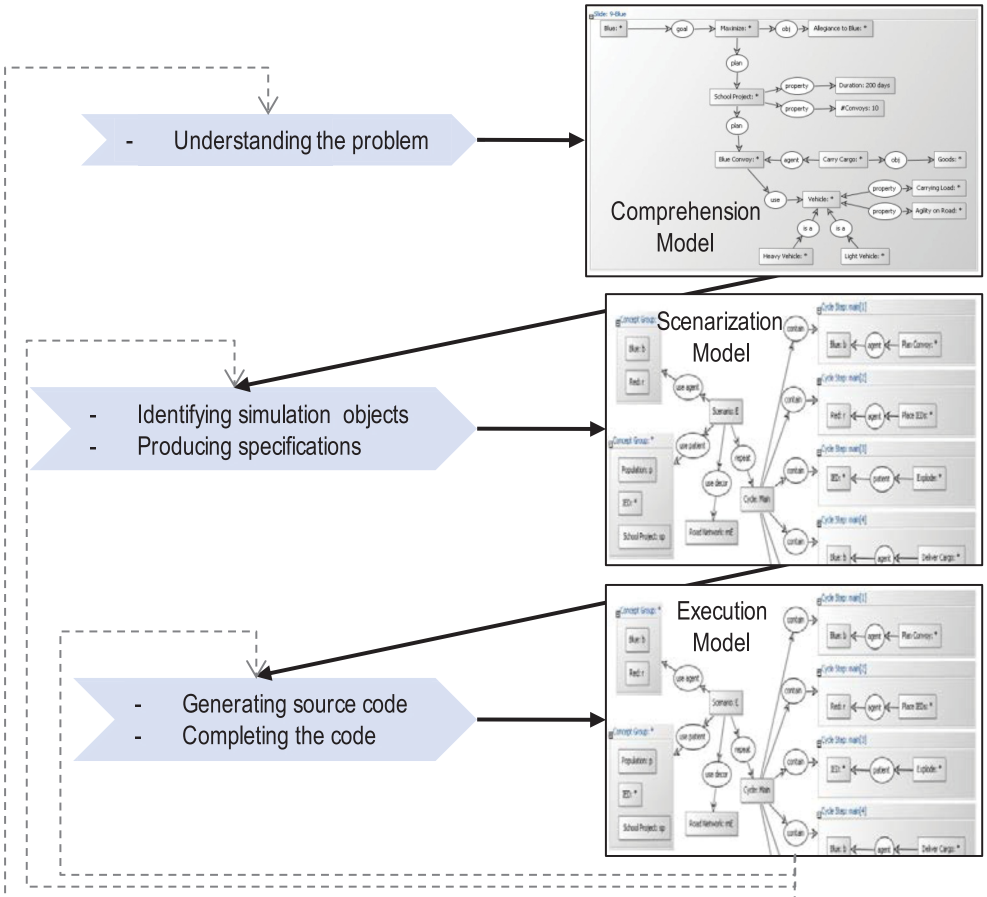


Figure 24: Running and revisiting the simulation.

The I-SCE concept feasibility has been demonstrated through a prototype part of the IMAGE V2 toolset. IMAGE-CoGUI V2 was used to create a scenarization model and ECLIPSE was used as the JAVA development environment to complete the execution model.

5.2 Related work

Agent-based modeling has been of interest to researchers for some time now. Some research has focused on the analysis and design of such software, but none has truly addressed the need for automated assistance in creating agent-based simulators from initial problem comprehension. This I-SCE approach is addressing this gap and supporting the spiral process of generating an agent-based simulator. In particular, it proposes an approach addressing the lack of formalism and procedural maturity stressed by Macal and North (2010). The generic Scenarization Vocabulary builds on previous work on planning systems e.g., STRIPS in Nilsson (1980) and SAIRVO in Lizotte and Moulin (1989).

In addition, the process of creating conceptual graphs to express a situation in the agent-based formalism can require a large amount of work, depending on the complexity of the situation. It allows structuring how one views a problem, and potentially helps in the comprehension process. A key step toward obtaining an executable simulation model consists in implementing the agents using the formalism of an agent-based simulator e.g., MASON (Luke, Cioffi-Revilla, Panait, Sullivan, & Balan, 2005), Ascape (Parker, 2012), Repast Simphony (North, Howe, Collier, & Vos, 2005). In the current work, the large amount of information contained in the conceptual graphs is exploited to automatically generate executable models rather than forcing the designer to do this work himself. This eliminates potential transcription errors between the conceptual and simulation models, keeps both models synchronized, and allows for their straightforward iterative development.

5.3 Scenarization vocabulary: Identifying simulation objects

The famous quote by Charles F. Kettering, a US electrical engineer and inventor (1876-1958), states: "A problem well stated is a problem half solved". This saying summarizes the intent behind the comprehension model which results from an individual or a team work in posing the problem at hand. Once well stated, the user needs to operationalize the situation. Starting from comprehension model concepts and relations, the user identifies simulation objects. Some comprehension elements will not be part of the Scenarization Model, either because they are not relevant to the simulation or they are not mature enough to be integrated at the current stage of the scenarization spiral process. Conversely, some elements will be introduced in the Scenarization Model to add details required for the actual execution of the simulation. In summary, the user isolates comprehension concepts and relations that are relevant for simulation and add concepts and relations required for simulation purposes.

As for the comprehension model, a set of graphs and the Scenarization Vocabulary constitute the Scenarization Model. All objects essential to simulate the situation needs to be consistent and defined at the correct level of detail. This is accomplished using both the Scenarization Vocabulary and the Specification Schemas (see next section). While the vocabulary is unconstrained for the development of a comprehension model, the usage of the predefined scenarization vocabulary is mandatory: Any concept or relation playing a role in the simulation have to be linked, through its type, under the generic scenarization vocabulary.

The main scenarization concept types are: Actor, Decor, Behaviour, Object Attribute and Predicate. Actors carry out behaviours that modify the situation. Such behaviours are triggered through actors' motivation or reflex rules. The situation includes all simulation objects along with their attributes. Decors are part of the situation but are never modified such as a road network that is never affected by the simulation but important for its execution. Object attributes are either variables (also referred as dynamic properties) that can be changed by the simulation execution or parameters (also referred as static properties). Predicates are logic statements detailing a concept such the IED Event concept described as a specific IED actor concept related to an Explode behaviour concept in Figure 11. The complete description of scenarization concept types is provided in (Lizotte & Rioux, 2010).

Figure 25 shows comprehension graph concepts linked under the generic Scenarization Vocabulary through their specific concept types. Generic types are part of the scenarization vocabulary while specific types come from the comprehension model. For each comprehension graph element, the user must determine whether or not this element will be part of the simulation and how it will fit into the generic Scenarization Vocabulary.

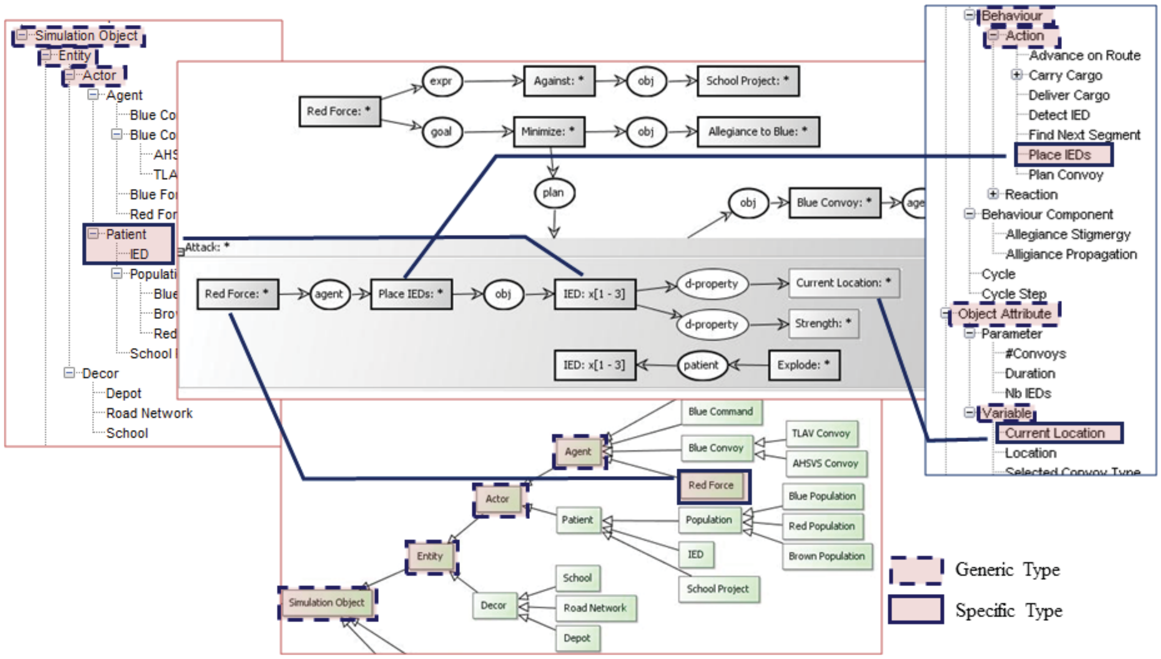


Figure 25: Linking comprehension concepts with the generic Scenarization Vocabulary.

5.4 Schemas: Working out the scenario *Mise-en-scène*

After a key simulation concept (such as an agent or an action) is identified, the user can start describing it using a specification schema. Specification schemas (also called prototypics) are associated with the following scenarization concept types: Agent, Patient, Decor, Behaviour and Scenario. The schema notion is defined in Chein and Mugnier (2009). In addition to specification schemas, construct schemas (also called patterns) facilitate the user's work. Construct schemas are not about adding relations with other concepts to a specific concept type, but rather about a nested graph pattern that can be used to detail a concept. For instance, a precondition to a behaviour can be detailed using the attribute predicate schema "Simulation Object:* → attribute → Object Attribute".

Figure 26 shows an example of such a graph "root" linking simulation elements according to the Scenario schema. It identifies agents, patients and a decor that are part of the scenario E, and specifies the different cycle step behaviours of the main cycle.

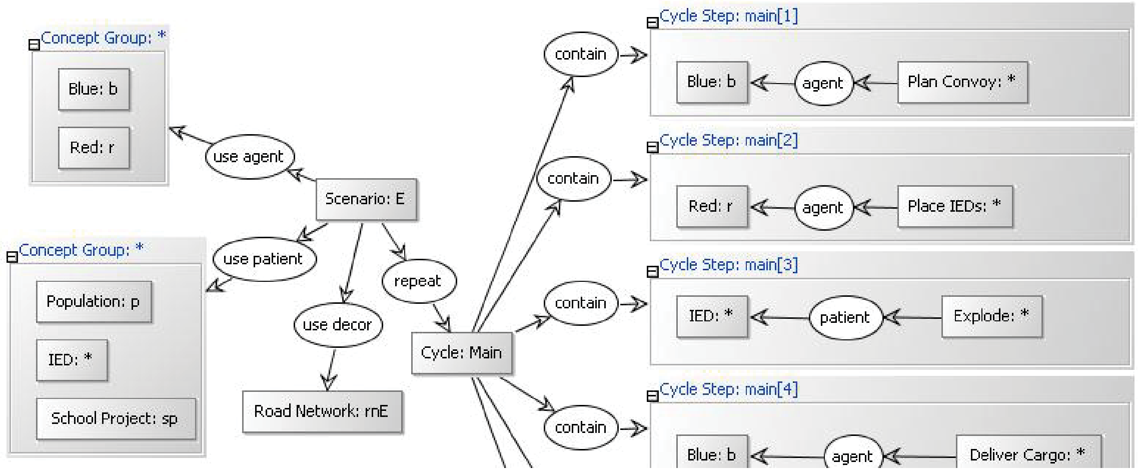


Figure 26: Specification of a scenario.

Figure 27 presents the patient schema and a graph resulting from its usage to specify an IED.

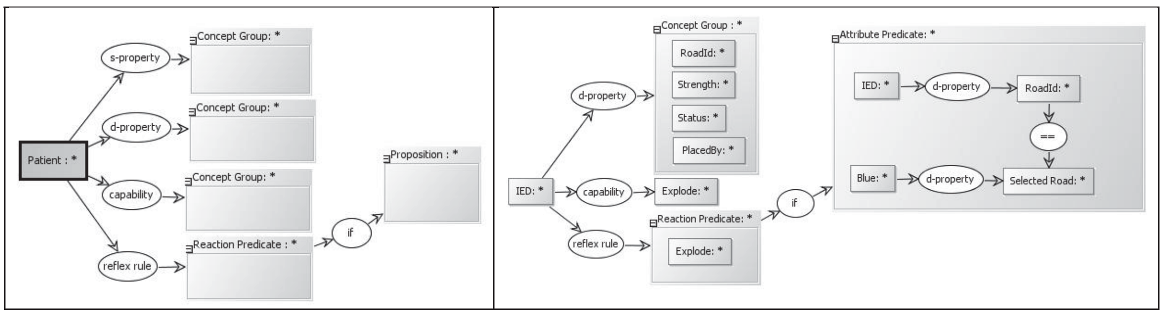


Figure 27: Specification of an actor.

A Behaviour Schema works in the same way but uses different relations such as precondition, co-condition and effect.

5.5 Java simulation framework: Completing the implementation of the simulation dynamics

The I-SCE approach would not be complete without its agent-based simulator framework. Use of this Java framework is the last activity a user must complete before actually running a simulation. From a root graph scenario, the user triggers the creation of the fundamental simulator software package. This transformation process uses the conceptual graphs (reachable from the root graph scenario) as input and produces implementation source code files specific to the problem. Added to the other files part of the framework, they constitute the fundamental simulator software package. The simulation engine of the framework is an adaptation of an interactive simulation software known as Multichronia (Rioux, Bernier, & Laurendeau, 2008a).

Most of the generated code is complete and does not need any additional work by the user, though some does. The user will find in this package a set of implementation classes, including method stubs, used mainly to detail relevant simulation dynamics. These are behaviour, stimulation and simulation cycle classes. More details are provided in (Lizotte & Rioux, 2010) and (Rioux & Lizotte, 2011). Although there are plans to improve graph notation and minimize code writing by the user, this stub approach was chosen deliberately for two main reasons. It allows a maximum of flexibility enabling any code insertion of interest to the user. In addition, the graph notation exists to ease the designer's work and, in many cases, behaviour is more easily expressed with a programming language rather than a graph notation.

5.6 Software architecture: Generating code using graphs

Three software components are involved in the generation and execution of a model. More details are provided in (Rioux & Lizotte, 2011). Firstly, the user works with the I-REP application. Within this Java application, one works out a scenarization model using the scenarization vocabulary introduced above. Using I-REP, the user can then trigger the transformation process during which scenarization graphs are converted into a conceptual intermediate model to ease further manipulation.

The second component, the scenario generator, is a Java program run within the Eclipse integrated development environment. It transforms the intermediate model conceptual graphs into an executable model. This Java transformation component parses and interprets the graphs, asks the user for input should there be missing values, and writes the result to appropriate data structures. These structures are then processed by two separate transformation data pipelines.

The first pipeline generates executable models including class descriptions defining generic classes as well as their attributes. Each class description is then provided as input to a code generation template and results in Java code including class skeletons and management code generated by the Apache Velocity toolkit (Apache Foundation, 2010). A template is available for each type of class file to be generated: static properties, dynamic properties, entity components, behaviours, motivations, reflexes and simulation cycles. One has to implement the dynamics of the model by filling in the skeleton classes of the cycles, behaviours and stimulations within the Eclipse environment. The resulting source code is then compiled to Java byte code, ready to be loaded by the simulator engine with the scenario name as an argument.

In the second pipeline, a data structure for the simulation engine framework is constructed. This structure contains all the information required to load the initial state of the generated scenario. This structure is marshalled to an XML file using the JAXB (Fialli & Vajjhala, 2005) marshaller implementation.

Once the scenario generation is complete, one proceeds with the simulation using a I-SIM tool. The generated simulator takes as input the XML scenario and dynamically loads the generated Java classes in order to form the initial simulation state. Then, the simulator scheduler executes simulation cycles until an end condition is reached.

6 Simulation: Testing the model

Computer simulation exploits approximate models of real situations (Santner, Williams, & Notz, 2003) for a variety of applications, including improving the understanding of CSs. Using computer simulation is often more powerful and cheaper than using a real system, for once the model is built and validated, it can be studied under a large variety of conditions almost freely. The possibilities of instrumenting are also greater. Computer simulation allows a person to confront his preconceptions and explore a CS thoroughly in a systematic and repetitive manner. It should then help to improve comprehension of CSs. The first section of this report situates the IMAGE Simulation (I-SIM) concept in regard of other IMAGE concepts. Is the simulation of CSs particular in some ways in regard to traditional simulation? Section 6.2 begins with an answer to that question by presenting a brief literature review of simulation-related techniques developed to tackle complexity. Then, a simulation tool called Multichronia for studying CSs interactively is presented. More details can be found in previously published reports and papers (Bernier, Laurendeau, Rioux, DeRainville, & Lizotte, 2010; Bernier, 2009; Rioux et al., 2008a; Rioux, Bernier, & Laurendeau, 2008b, 2008c; Rioux, Laurendeau, & Bernier, 2010; Rioux, 2008a, 2008b).

6.1 The simulation concept

The I-SIM concept is situated between the scenarization concept (I-SCE) and the exploration concept (I-EXP) as depicted in Figure 28. I-SCE module generates an executable model based on the representation model developed by the user and on its mapping in an executable format. This executable format can be built from scratch or resulting from the customization of an existing simulation framework. I-SIM does not implement the executable code itself but only controls the parameters values, execution, and branching. I-SIM also provides functionalities for studying the CSs under various conditions. Finally, I-SIM forwards (via a database or an extensible mark-up language – XML – data stream) simulation data to the I-EXP module to be visualized by the users. Along with the data transmitted, I-SIM includes information (meta-data) that contextualizes in which conditions the simulation was done (e.g. the experimental design).

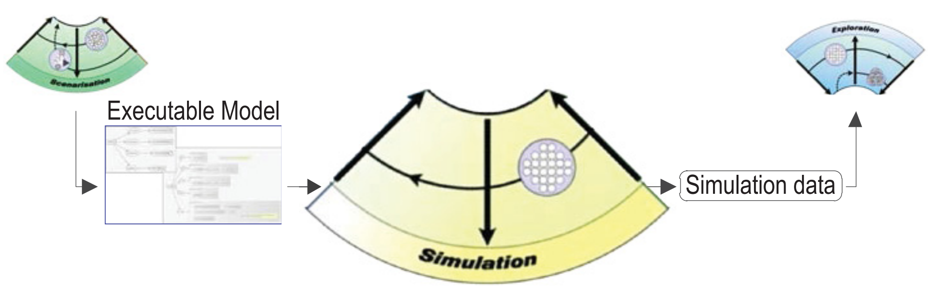


Figure 28: The I-SIM concept.

6.2 Related work

Although complexity is often mentioned in application-specific papers involving simulation, few papers explored and elaborated theory at the intersection of complexity and simulation. Edmonds (2005) is one of them and he proposed a syntactic and a semantic complexity. The first one relates to the computational distance between the input (initial conditions) and the outcome of a simulation. For this type of complexity, there is no shortcut: a lot of computation is required and the simulation is usually characterized by many variables for which the level of importance, i.e. the effect on the outcome, is not known. Climate changes and meteorology models are examples of syntactic complexity. Some simulation techniques presented in the next paragraph tackle this kind of complexity. As for the semantic complexity, it refers to the difficulty of capturing a real system in a formalized way. I-REP helps to deal with semantic complexity and, in combination with the simulation module that overcomes the syntactic complexity, at instantiating the current comprehension model so the user may compare his current mental model with the outcome of the simulation. Such computational-mental model loop allows learning from evidences with simulated CSs (Sterman, 2000). This feedback process is rooted in learning theory (Argyris & Schön, 1978) which exploits single-loop and double-loop learning. Such learning approach is based on actively testing the model and from observing the response which may force the learner to adjust his abstract model to reflect the simulated reality. In the context of IMAGE, representation models of CSs are created dynamically by subjects matter expects in I-REP. These models can be simulated only once a rigorous process had eliminated incoherencies and errors, similarly to the role a compiler before executing a program. The consequence is an increased rigour and clarity. Then, simulations can be performed to confront these mental models and their explicit representation models with simulation outputs. Mental or representation models can then be updated to resolve contradictions. The repeated application of this process should bring users to converge toward a common, consistent and comprehensive representation and computational model. This approach depends on model validity (Law, 2007) and users should not interpret more from these models than they can provide.

Many existing simulations approaches, little to no interactive, can help tackling complexity. Some designs of experiment (Montgomery, 2008), like fractional factorials, can be exploited when facing CSs that are characterized by a large number of factors. However, if the number of factors is very large, other approaches like nearly-orthogonal Latin hypercube (Cioppa, 2002; De Rainville, Gagné, Teytaud, & Laurendeau, 2009) are more relevant. According to the number of input variables studied (factors) and the complexity of the response surface (output), specific experimental designs may be selected (Kleijnen, Sanchez, Lucas, & Cioppa, 2005). If computational time is an issue, meta-models (Kleijnen et al., 2005; Kleijnen, 1992) that approximate the input-output relationship can substitute to the original model. This method is not suited to study CSs that are characterized by non-continuities and by rare but heavily consequential events though. Instead, data farming (Brandstein & Horne, 1998) and agent-based distillation (Horne & Meyer, 2004) exploits vast computer resources to study computer models under billion of conditions so both general tendencies and rare events (outliers) may be identified. Exploratory analysis is another approach that uses models to get insight into a CS. Massive scenario generation (Davis, Bankes, & Egner, 2007), exploratory analysis (Davis, 2000) and emergent behaviour exploration (Gore, Reynolds Jr., Tang, & Brogan, 2007) are examples of semi-interactive applications of this approach. What if analysis (Golfarelli, Rizzi, & Proli, 2006; Philippakis, 1988) is another approach, if carefully exploited, can improved CS understanding and decision-making. These methods are useful but lack the interactivity required to support a team of subject matter expert engaged in a highly iterative hypothesis formulation and testing approach such as described in Chapter 2.

Sterman (2000) idea of learning form evidences using action-observation loop within a simulation environment is close to what IMAGE tries to achieve in term of interactivity, especially for revealing counter-intuitive phenomena, unveiling evidences concealed by complexity, and considering feedbacks. Many fields explored similar ideas. Interactive simulation (Rothrock & Narayanan, 2011), which consists in dynamically influencing the outcome of a computer simulation like in a computer game or a flight simulator, is one of them. Interactive simulation was also embodied in visual interactive simulation (VIS) (Hurrion, 1976) and other similar approaches (Standridge, Matwiczak, Davis, Musselman, & Brunner, 1990; Wright, 2004) that are well suited for analysing complex and dynamic systems. Experiment with VIS (Bell & O'Keefe, 1987) showed that, when users try to solve resource allocation, they performed badly compared to formal experimentation but performed well when compared to a priori solution. Computational steering (McCormic, Defanti, & Brown, 1987) is a specific type of interactive simulation where a user may intervene in autonomous computational processes, usually those requiring super-computation. As stated by McCormic et al (1987), "scientists want to steer calculations in close-to-real-time… change parameters… and see effect". Visual interactive simulation was retained as the most promising approach but is not enough formal and generic. Multichronia was proposed to fill that gap.

6.3 Multichronia

Multichronia was developed as the I-SIM tool for IMAGE, the tool being identical for V1 and V2. Multichronia is a generic VIS framework (Rioux et al., 2008a) that supports both what-if analysis and formal experiment analysis. It offers visual and interactive means for managing many concurrent executing simulations. It supports parameter, simulation, data and visual space exploration as illustrated in Figure 29. It should be noted that the visual space exploration is implemented by the I-EXP module in IMAGE.

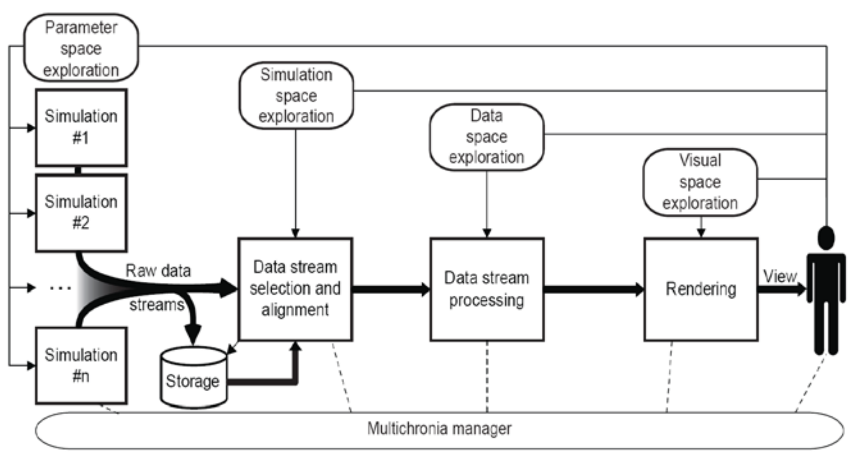
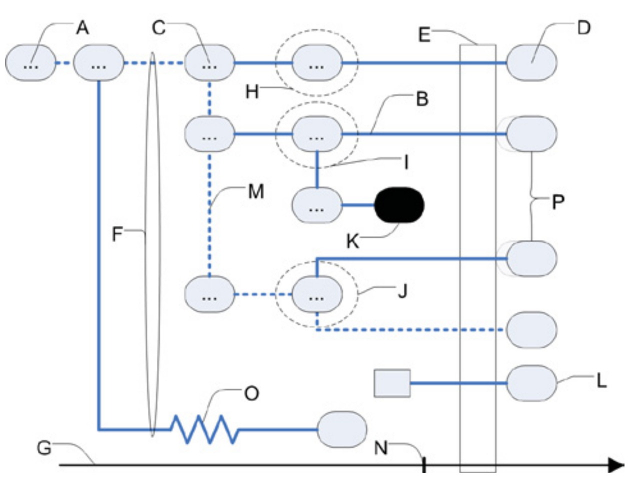


Figure 29: Multichronia – Conceptual framework.

Multichronia comprises a multichronic tree for providing a standardised visual representation of all simulations and activities, a set of activities for supporting simulation exploration and a generic data pipeline for streaming results to the I-EXP module.

6.4 Multichronic tree and activities

Visual representation activities performed on simulations has been proposed by HyperScribe (Brodlie et al., 1993), TRICEPS (Pickles, Haines, Pinning, & Porter, 2004) and Fischer et al. (2007) by exploiting historic trees of multiple simulations. These tools lack of support for many activities envisioned by I-SIM. A new approach was then required. Multichronia implements a multichronic tree that defines the elements that represent visually activities performed on a simulation set (see Figure 30).



A. Tree root: initial scenario (initial conditions)
B. Tree branch: partial run
C. Tree node: divergence point
D. Tree leaf: simulation instance
E. Parallel segments: window of comparison
F. Parallel branches: aligned points
G. Horizontal axis: axis of reference
H. Simulation parameter modification (simple)
I. Creation of a new divergence point
J. simulation parameters modifications (multi.)
K. Terminated simulation
L. Pruned simulation: insignificant simulation
M. Path from the root to the leaf: run
N. Location on the axis: point
O. Zigzag sign: continuous parameter change
P. Moving nodes: simulation playbacks

Figure 30: Multichronic tree depicted with its elements and definitions.

In addition to the multichronic tree, Multichronia provides a rich set of interactive activities to support IMAGE concept: create and destroy a divergence point, simulate and playback several runs simultaneously, pause / resume / stop a simulation playback, go forward and backward in time, vary the simulation playback speed, change the variable of reference, synchronize multiple simulation playbacks, change the layout of the tree, etc. Multichronia is not a simulator; it interfaces with existing simulator but under some conditions (Rioux et al., 2010). Failure to comply with these conditions may reduce significantly the number activities available. If executable models are created from the IMAGE representation tool, then Multichronia functionalities are all available. Figure 31 shows the Multichronia graphical user interface programmed in JAVA by using the Processing library (http://www.processing.org). Figure 32 shows Multichronia tailored for the experimental set-up.

6.5 Data pipeline

The Multichronia data pipeline (Rioux et al., 2008c) uses an XML based-approach to implement a generic data pipeline for a variety of simulation and visualization (exploration) tools. The XML pipeline exploits Java architecture for XML binding (JAXB), XQuery, extensible style sheet language transformations (XSLT) and native XML database and takes its roots in model-driven engineering (Kent, 2002), for it allows a platform-independent modeling and an automatic implementation on specific platforms.

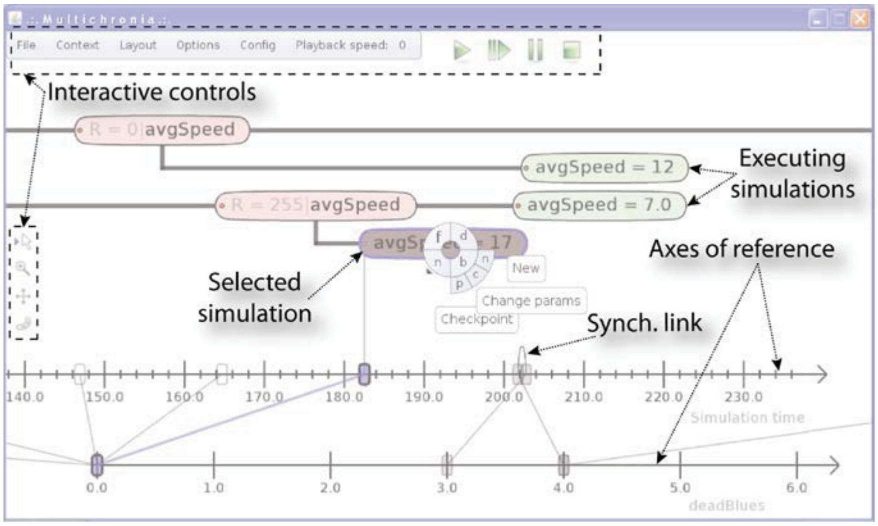


Figure 31: Multichronia graphical user interface.

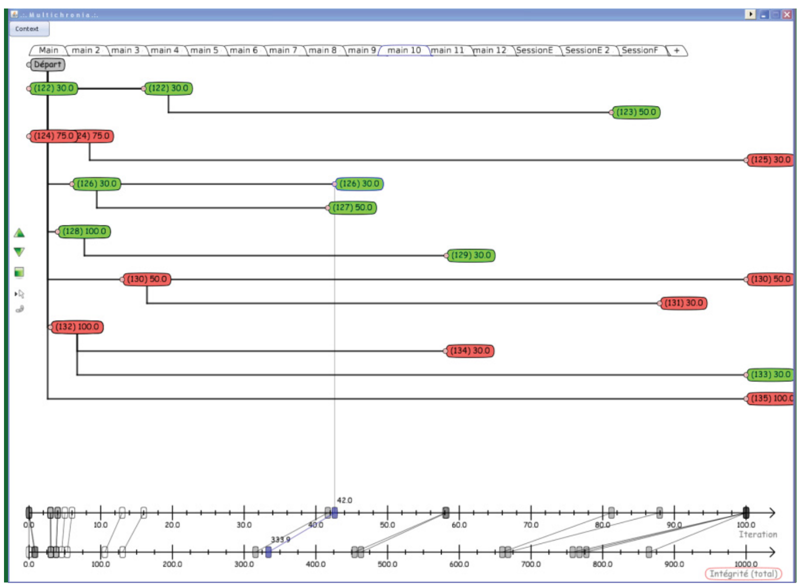


Figure 32: Experimental version of Multichronia.

6.6 Interactive design of experiment

A prototype of an interactive design of experiment (IDoE) tool was developed within Multichronia. IDoE aims at providing usual computer simulation experiment tools to study CS under many conditions by providing sampling, optimization and interpolation algorithms (see Table 1). IDoE was excluded of the experiment because it would have provided direct answers to all questions at the tactical level for participants of IMAGE condition.

Table 1: IDoE functionalities provided by IMAGE.

Sampling	Optimization	Interpolation
Scrambled Halton (Bhat, 2003)	CMA-ES (Hansen & Ostermeier, 2001)	Support vector regression machines (Smola & Vapnik, 1997)
Nearly orthogonal Latin hypercube (Cioppa, 2002)	Simplex (Murty, 1983)	Kriging (Cressie, 1990)
Full-factorial (Montgomery, 2008)	SPSA (Spall, 1998)	

IDoE is implemented such that the interactivity with the user is maintained and continuous visual feedback of the background simulation progress is provided (see Figure 33). Algorithms are executed in background on a multiprocessor or remotely so that results be obtained faster.

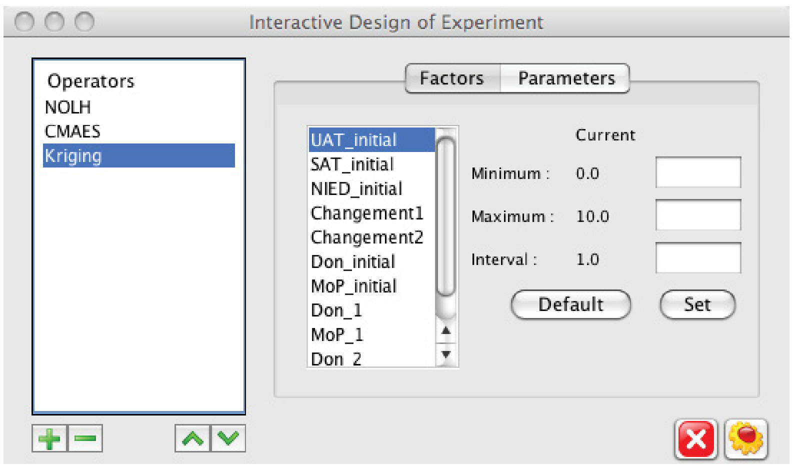


Figure 33: Interactive design of experiment graphical user interface.

6.7 Discussion

I-SIM is implemented via Multichronia that provides a multichronic tree, a set of functionalities for interacting with simulation runs and a data pipeline for transmitting results to the I-EXP module. Cost and time constraints allowed implementing a fraction of the desired functionalities. Missing functionalities include a more sophisticated IDoE, more information (e.g. execution progress) displayed on the multichronic tree, a more user-friendly interface (e.g. Tableau like interface with drag and drop of parameters on the branches), copy and paste of changes between branches, etc. As it will be seen later, the fact that Multichronia was the preferred and the most used tool in IMAGE during the experiment as well as that it was the only for which the time spent to use it was correlated with comprehension performance is an encouraging element to pursue its development.

7 Exploration: Making sense of large datasets

In a wide range of scientific and technological domains (e.g. economics, climate and seismic modeling, melting modeling, astronomy, neuroscience and archaeology), experts need to make sense of and extract useful knowledge from large datasets composed of various types of data. To achieve these goals, experts need tools combining visualization and rich user interactions to guide them to incrementally and interactively explore large datasets, to organize data, to process information, and to go through and understand these datasets. The IMAGE Exploration (I-EXP) module consists in making datasets (generated by the I-SIM module) clear in tailored views, which can be used by users to augment their individual or collective comprehension models.

7.1 The Exploration concept

Figure 34 illustrates the I-EXP concept which aims to produce individual or collective explanatory and tailored (static and / or dynamic) views in ways that words cannot communicate clearly. These views, in order to bring a meaning to (raw and / or processed) data, have to generate, stimulate, increase or accelerate the CS understanding. The concept is supported by the I-EXP tools, which implement visualization and interaction techniques and / or mechanisms by, taking advantage of the potential of different platforms in terms of technology and in terms of exploitation of human sensory information (mainly *perception* and *action*). The I-EXP concept has been developed through two main prototypes: the I-EXP V1 composed of I-EXP V1 Tactical (at the low level of complexity) and I-EXP V1 SOP (at the high level of complexity), and the I-EXP V2. I-EXP V1 was part of the IMAGE V1 cognitive experiment focussing on individual understanding while I-EXP V2 was part of the IMAGE V2 feasibility effort and focused on immersive virtual exploration and collaboration aspects.

As represented in Figure 34, hardware components composing the platforms supporting the I-EXP tools span from a traditional hardware setup composed of LCD screen / keyboard / mouse to a more high-tech platform exploiting immersive technologies leading to the user's immersion in a virtual environment (VE). In fact, the I-EXP module puts emphasis on the proper use of technologies in order to fulfill the user's exploration needs: from (1) conventional non-immersive systems (no immersion, limited interaction), to (2) hybrid systems combining the adequate use of non-immersive systems and immersive interactive systems (in terms of head-centered interaction, access to traditional interaction tools augmented by virtual gadgets) up to (3) high-end fully immersive systems (in terms of large displays, large field of view, head- / hands-centered interaction, more natural interactions). Immersion can be achieved to varying degrees by using advanced human machine interfaces (involving sensory modalities – visual, auditory, tactile, olfactory – and body tracking, etc.), and by exploiting more natural interaction with the immersive VE (IVE). Immersion may lead to a sense of presence which is a psychological emergent property of an immersive system (Sadowski & Stanney, 2002). This property refers to the user's sense of "being there" which means that the user feels he is part of the environment because he forgets that his perceptions are mediated by technologies. Immersion is a *necessary* rather than a *sufficient* condition for presence (Youngblut, 2006).

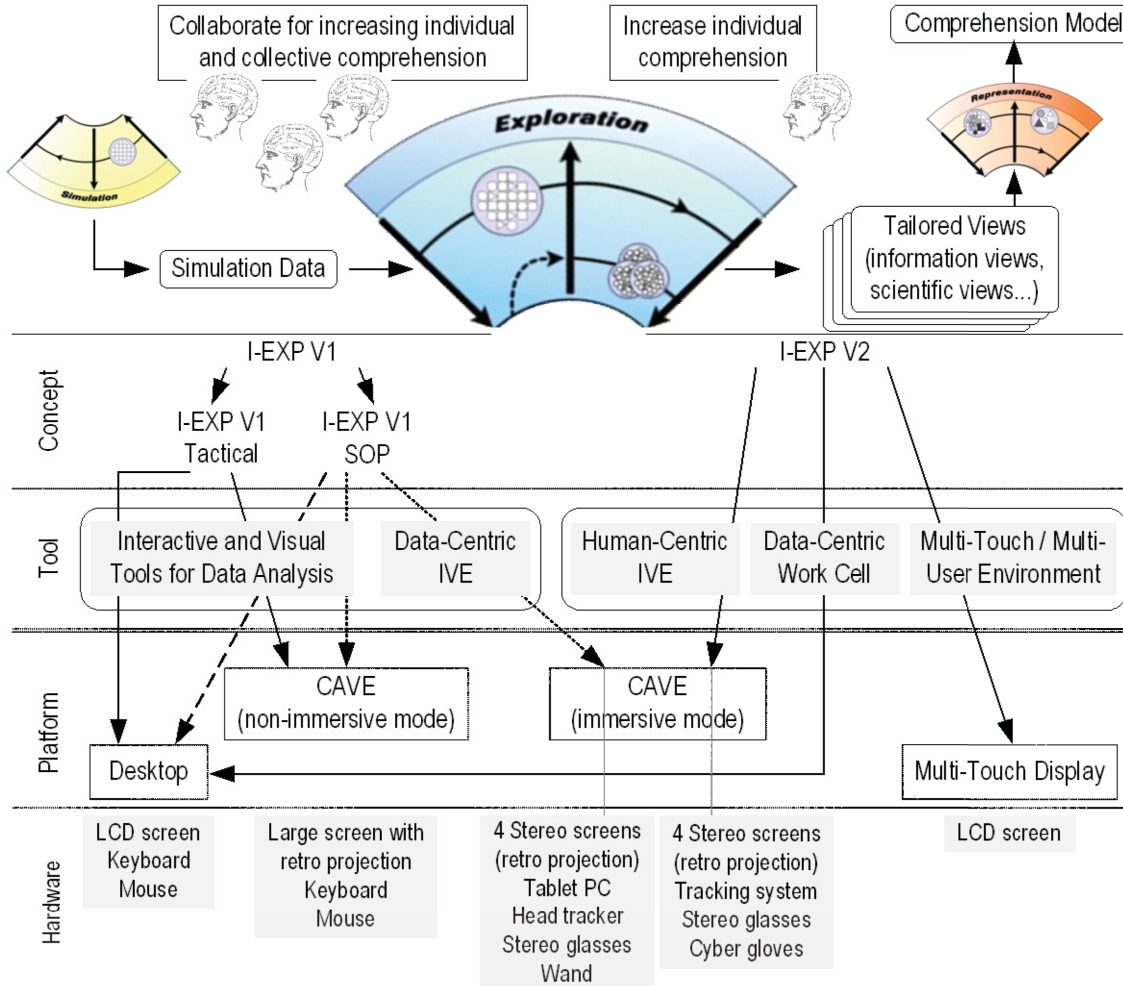


Figure 34: The I-EXP module at a glance.

Virtual reality (VR) is a technology that immerses one or more users in a VE representing a real or fictive world that allows users to become agents capable of changing the properties of the environment and interact with its entities (Burdea & Coiffet, 2003). Since the advent of VR technology, researchers were particularly interested in the 3D interaction which can be regarded as the driving component behind any interactive system. Indeed, interaction gives a better feeling of immersion and plays a crucial role in the determination of presence. IVEs, such as CAVE (Creagh, 2003; Cruz-Neira, Sandin, DeFanti, Kenyon, & Hart, 1992) support, among other things, the ability to analyze and interpret a large amount of data faster and to explore data in a more intuitive way through interactions with visually rendered data. The CAVE (from Mechdyne, http://www.mechdyne.com/) used in IMAGE is a multi-screen projection-based VR system with three walls and a floor that are arranged in a cube for total immersion. One of the major advantages of the CAVE is that it makes (collaborative) exploration of large-scale content possible through manipulation and navigation within an IVE. The user navigates the virtual content by naturally moving around inside the cube, while his field of view is filled with 3D representations of components of the environment. The CAVE is viewed as a spatially immersive display because it enables simultaneous access to a huge amount of data more directly than any other display devices and helps escaping the conventional bias towards 2D computing by organizing content more effectively in 3D. To achieve realistic interactions with immersive displays, the tracking system (IS-900 system from Intersense, http://www.intersense.com) implements fast refresh rates, low latency and smooth tracking. Consequently, it provides smooth and precise estimates of the position and orientation (6-DOF) of the user's head / user's hands while not interfering with the user's immersive experience. The tracking system offers ergonomically designed devices: (1) a head tracker attached to stereo glasses (CrystalEyes®3 from StereoGraphics Corp., http://www.reald.com); (2) trackers attached to Cybergloves (from Immersion Inc., http://www.cyberglovesystems.com) which are composed of 22 flexible sensors that can accurately measure the position and movement of the fingers and wrist; and (3) a tracked wand (from InterSense) which includes buttons (four on one side and a trigger on the other side) and a joystick for interacting with the virtual content. Stereoscopic rendering based on active goggles provides the user with depth perception of the virtual content.

7.2 I-EXP V1

This version implements interactive and informative simulation analysis tools for exploring different aspects of complex scenarios (a more extended and thorough study and more details on some I-EXP V1 tools are presented in (Mokhtari & Boivin, 2013a)). The tools developed in this first version mainly offer the user rapid prototyping of views, as well as tailored views and focus on the "best view" concept rather than the "perfect view" concept. Furthermore, to address as much as possible exploration needs with respect to diversity and richness, the toolset proposed is a combination of turn-key and commercial off-the-shelf (COTS) tools. In the spirit of the *Visualization* domain, the main task of which is to allow *information* to be derived from *data*, the I-EXP V1 toolset has been thought to assist experts / users understand and interpret large and rich datasets, by exploiting information rendered graphically (e.g. graphs, diagrams, charts, etc.) which are visual artefacts representing data, illustrating relationships and patterns among data, presenting *information* (appropriate conversion of *data*), and also allowing analysis and comparisons between graphical representation of datasets. In addition, the I-EXP V1 toolset also addresses the dynamic display of trends in the data and human interaction with the displayed data. Consequently the toolset has been designed to: (1) explore and compare large datasets; (2) extract (hidden) information; (3) discover unexpected trends in the data or find patterns; and, finally, (4) understand specific non-intuitive aspects of CSs. To achieve these goals, the I-EXP module uses different techniques such as (1) filtering (for focusing on relevant subsets of data) and data brushing (support visual linking of various data and address the visual fragmentation problem of multivariate data representations) (Becker & Cleveland, 1987), (2) datasets comparison, (3) multi-level datasets exploration (at Tactical and SOP levels), and (4) real-time and stereoscopic rendering.

For this version, a key architectural point to consider is the synchronization between the I-SIM and I-EXP modules. This synchronization means that the selection of an item in the I-SIM module (such as a simulation) must be reflected in the views developed in the I-EXP module according to a visual paradigm that highlights focused data. Multichronia then serves as the primary control point for user. The I-EXP V1 toolset can be used on a traditional desktop and on a large screen (see Figure 38) as well as on a display wall and in a CAVE, both providing active stereoscopic visualization (see Figure 39).

7.2.1 Architecture

The exploitation of complementary exploration tools is at the core of the I-EXP module, increasing the number of analysis opportunities available to the user. In addition, many COTS technologies were included in the I-EXP module.

The data exchange between the I-SIM and I-EXP modules is a difficult problem to address because of multiple XML data streams that need to be transferred in parallel from the I-SIM module to the I-EXP module. The usage of a database[3] (DB) to create data persistence and centralisation has been adopted. To avoid the I-EXP module to consult the DB too frequently, the I-SIM module notifies the I-EXP module of changes occurring in real-time such as creation, deletion or selection of one simulation. For that purpose, a synchronization application (communication channel) has been created between the I-SIM and I-EXP modules. Managing multiple connections and executing data processing in parallel are performed by using the ACE framework (http://www.cs.wustl.edu/~schmidt/ACE.html).

The analysis toolset (at the Tactical and SOP levels) has been developed with the Eye-Sys software package (http://www.eye-sys.com/), which is a commercial visualization application. It gathers data from various sources adopting different formats, manipulates them in real-time, and uses them to drive interactive and real-time visualizations. Eye-Sys proposes a generic and flexible platform that allows real-time visualization and interaction, data fusion with flexible input options, interoperability with other applications in real-time and comprehensive software development kit (SDK) and scripting. Furthermore, it is possible to create tailored modules since COM interfaces allow external programs to connect to these modules and control them. The Tableau software package (http://www.tableausoftware.com/), a visual analytics tool, was included in the I-EXP V1 toolset to explore data in a different but complementary way. Its strengths are mainly its user-friendly interface and the ease of learning and use. It exhibits several powerful functionalities, i.e. a “drag and drop” technique that is very intuitive and accessible to almost all interface elements and a support for (small to large) data sets in a wide variety of formats. These latter ones are handled automatically, quickly displayed in one or more dimensions, and represented using visual techniques such as color, shape, or size of the dots. Techniques such as zooming and filtering, combination of parameters, summation, and averaging complement Tableau’s analysis capabilities.

Figure 35 summarizes the main structure of the proposed architecture. The technical memorandum *IMAGE Exploration Module - Architecture* to be published in 2013 will report a more complete description of the architecture implemented for the I-EXP V1 toolset.

[3] The DB is a MySQL database (http://www.mysql.com/) managed with MySQL++ (http://tangentsoft.net/mysql++/) which is a library of C++ classes.

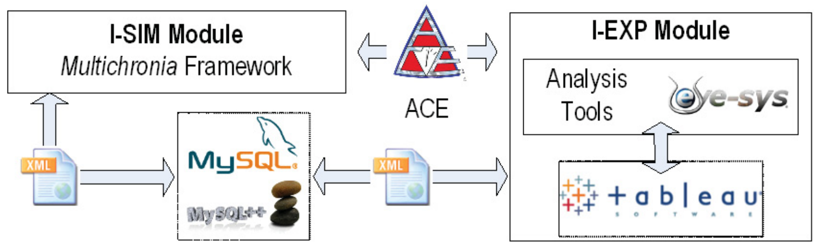


Figure 35: Software architecture of the I-EXP module for IMAGE V1.

7.2.2 I-EXP V1 Tactical tool (for a low level of complexity)

A simulation analysis tool has been designed at this level. On one hand, the tool allows the user to visualize, analyze and explore simulation data but also to understand how the whole set of parameters evolves during the simulation. On the other hand, the tool maps user-created simulations into simulation space (a kind of log file) composed of 726 combinations of variables. The tool's interface (visible on Figure 36) is mainly built around of five displays or panels of information (zones A to E) in which the user can visualize or interact with the simulation parameters. The simulation analysis tool allows the study of only one simulation at a time while Tableau (access to by zone E) performs the analysis (and / or comparison) of a set of simulation data. The main display (zone A) of the interface is a geospatial view in which simulation animation is performed, important events are highlighted, and information panels linked to each simulation agent (LAV, IED and RPG) are displayed on-demand. The user can change the camera viewpoint (zone F) at any time. Information panels (zone C) display the value of simulation input data (i.e. allocated blue forces – BF – and red forces – RF – resources) and output data (i.e. set of MoPs – integrity, timeliness and cargo). Furthermore, the user can see at a glance on one plot (zone B) the evolution of each simulation parameter. Finally, the spatial distribution of simulations (zone D) according to the (BF and RF) resource allocation (UAT, SAT and IEDs / RPGs) is available and updated at each creation or deletion of one simulation. The user always knows which area of simulation space has been visited or is under study and can steer his analysis towards the subspace he wants to investigate. Finally, the user can zoom, rotate, pan, and change camera angle in zones A and D. A more detailed description can be found in (Tye-Gingras, 2011).

7.2.3 I-EXP V1 SOP tool (for a high level of complexity)

Two complementary tools for mission analysis have been designed at this level: the mission data analysis tool itself and the parameters and MoPs analysis tool. The interface of the mission data analysis tool (visible at the top of Figure 37) is divided into two distinct zones: the information zone (on the left, zones 1 to 6) and the analysis zone (on the right, zones 7 to 9) in which the user can explore and analyse data as well: (1) Information zone – Zone 1: *MissionID* corresponds to the identification number of the current mission (mission focused by user) and *IterationID*, the identification number of the current iteration (Tactical simulation) which is one of the 100 iterations composing the mission; Zone 2: At any time, the user can access Tableau Software to analyze the current mission data or a set of mission data; Zones 3 to 6 correspond to information panels: zone 3 illustrates input data of an iteration for the whole current mission (at each iteration, the priority given by the BF evolutionary algorithm in terms of optimizing integrity or cargo is displayed with *color* cues while the percentage of cargo given to the population is displayed with *line* cues; zone 4 displays, for the current mission, the coevolution parameter performance, i.e. population allegiance, intelligence gathering, and support to BF (*line* aspect); zone 5 presents, for all missions created by the user, mission MoPs (cumulative iteration MoPs – integrity, cargo and timeliness). The current mission is highlighted as well (*chart* rendering); and zone 6 displays input and output data for Tactical simulation *IterationID* (an adapted information panel already available at the Tactical level, Figure 36 zone C); and (2) Analysis zone – Zones 7 to 9 are interactive panels exhibiting detailed and tailored display modules. Zone 7 allows the user to choose what to display and how it should be displayed. Zones 8 and 9 plot 2D or $2D^{1/2}$ graphics.

The mission parameters analysis tool (the interface is visible at the bottom of Figure 37) allows the user to create tailored views to better assess correlation between parameters. The main display (zone D) shows, on-demand, all available combinations of parameters: input data (BF resources – UAT and SAT; RF resources – RPGs and IEDs), MoPs (cargo, integrity and timeliness), co-evolution parameters (population allegiance, intelligence gathering, BF support, the cargo offered to the population), and the iteration priority (integrity or cargo) (*form* or *colour* aspect). The user can choose (via zone A) what mission or set of missions he wishes to explore. Each mission, composed of 100 iterations, is displayed as a chain of *coloured* and *scaled* spheres (representing one or several iterations). Depending on the parameters selected, all 100 iterations are not necessarily represented by a distinct sphere; some of them are lumped in a common sphere. The graphics is updated automatically to adapt to changes of parameter values. Three to five dimensions are always available to the user (zone B). The displays are updated dynamically each time a new mission is created or an existing mission is deleted in the I-SIM module, and when the user focus changes (current mission). The current mission is highlighted by a thicker chain, the current iteration is highlighted by a cross and an information panel is attached to it (zone E). The *color*, *form* and *size* codes are given by the corresponding legend (zones B and C). At any time, the user can zoom, rotate, pan, and change camera angle. More details can be found in (Girardin, 2012; Mokhtari, Boivin, Laurendeau, & Girardin, 2010).

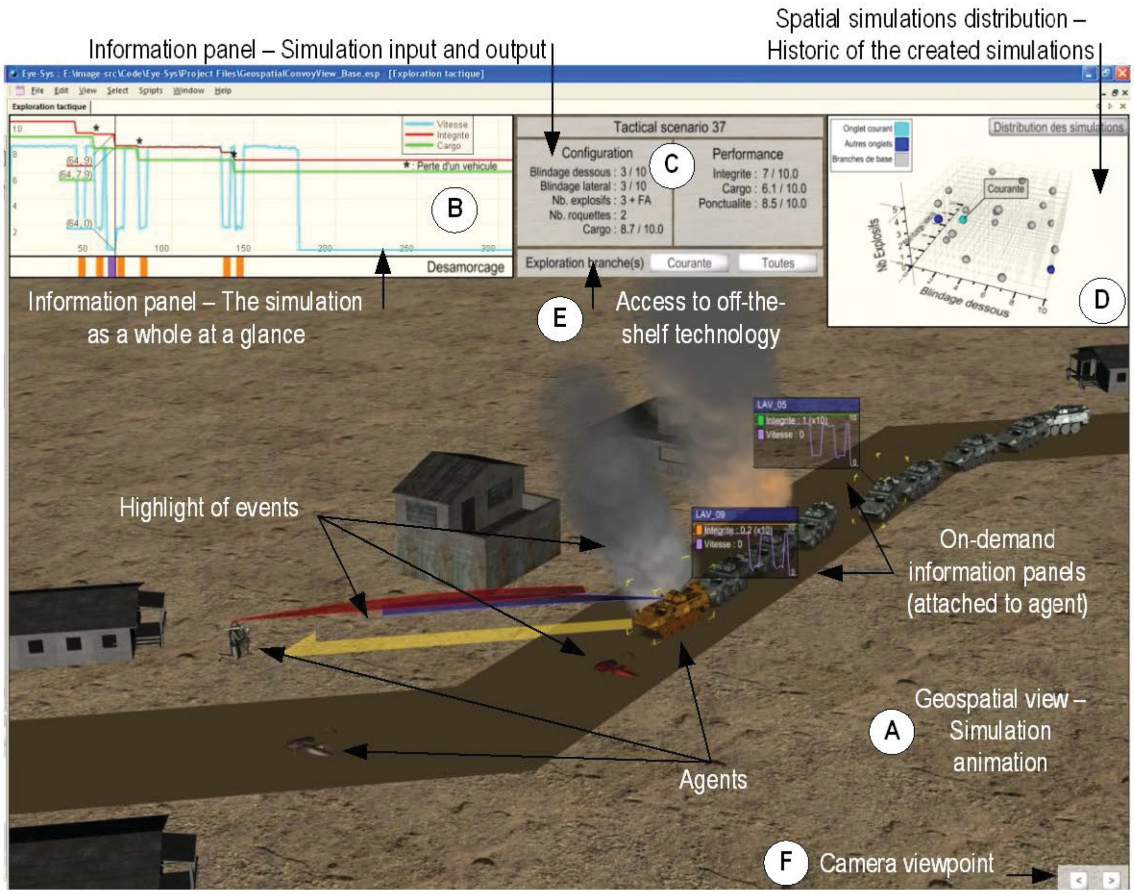


Figure 36: I-EXP V1 - Simulation analysis interface at the Tactical level.

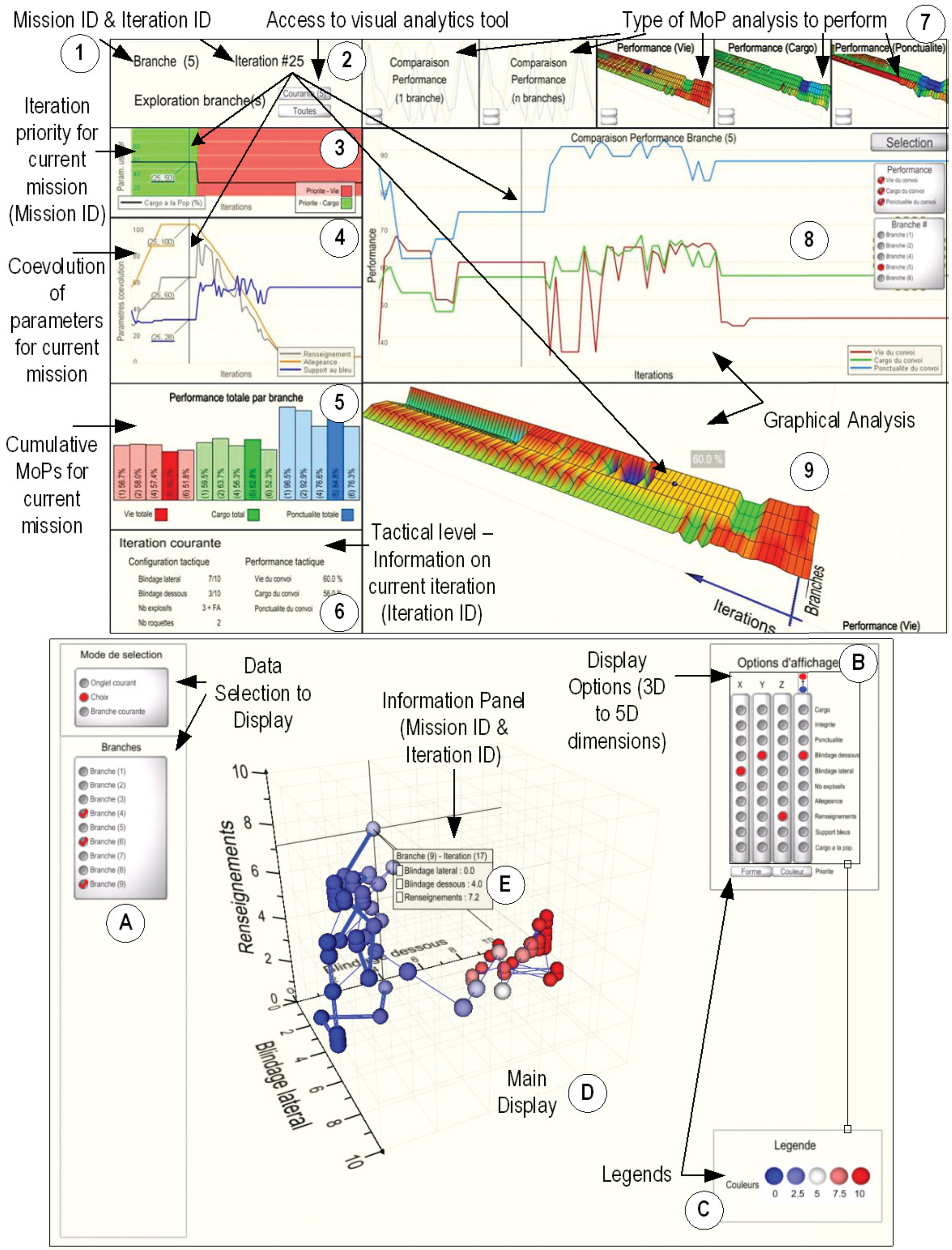


Figure 37: I-EXP V1 - Analysis interfaces at the SOP level: (top) for mission(s), (down) for parameters.

Figure 38: I-EXP V1 - Platforms used at the Tactical and SOP levels: (left) desktop, (middle and right) CAVE in the non-immersive mode.

CAVE adaptation – A data-centric immersive virtual environment

Although traditional techniques of data visualization remain conceptually interesting and viable when the user faces a (very) large amount of data to be processed, some techniques become more difficult to implement on a traditional desktop screen. VR allows for the ability to simultaneously analyze a large amount of data / information as well as for a more intuitive exploration in IVEs. The VR gives the impression to the user that he interacts and manipulates data directly and not just its graphical representation (Bryson, 1996). Desktop tools and parts of desktop tools have been adapted with this in mind so that the user can visualize and explore the data as being “in-the-box / inside looking out” rather than as being “out-of-the-box / outside looking in” (desktop version). The quality of a VR application depends on a multitude of factors, mainly the realism of the 3D environment, the fluidity of the graphics display, the sensory feedback and intuitiveness of navigation (Sherman & Craig, 2003)[4].

IVEs provide many benefits for understanding complex systems (Knight & Munro, 1999), namely the visualization of immersive 3D graphics and different immersion paradigms offering intuitive navigation to explore and analyze data / information. Besides, IVEs facilitate the learning process (Dalgarno & Lee, 2010) by, among other factors, promoting the spatial representation of knowledge and augmenting both user’s motivation and commitment.

Two data-centric IVEs have been developed; the mission parameters and MoPs analysis tool has been adapted to become the *Cube View* tool and the analysis and comparison graphics of missions in terms of MoPs have been adjusted to become the *Ribbon View* tool (Mokhtari, Boivin, & Laurendeau, 2013). The *Cube View* (on the left of Figure 39) allows users to visualize datasets and manipulate up to five parameters at a time (the three axes and two aspects to choose between *color*, *form* and *size*) with the same approach as the one defined for the desktop version. The *Ribbon View* (on the right of Figure 39) supports simultaneously visualization of the three MoPs throughout the 100 iterations of the SOP mission, and for all simulated missions. A tablet PC is used to switch between immersive and non-immersive modes, and also to select parameters and data to configure and display. The user can navigate inside the 3D virtual content at will. He uses the wand to manipulate (translate, rotate, and resize) the 3D graphics. Data selection is not needed because any manipulation involves the whole view. Figure 39 shows a user manipulating datasets with the wand and visualising them with the stereo glasses. The data-centric IVEs are supported by a custom open architecture, mainly based on openGL (http://www.opengl.org) and the 3D graphics toolkit OpenSceneGraph (OSG), a 3D graphics rendering engine, frequently used to

[4] To understand some concepts related to VR, the reader can refer to (Drolet, 2008), chapter 1.

develop applications in scientific visualization, VR and modeling (http://www.openscenegraph.org/).

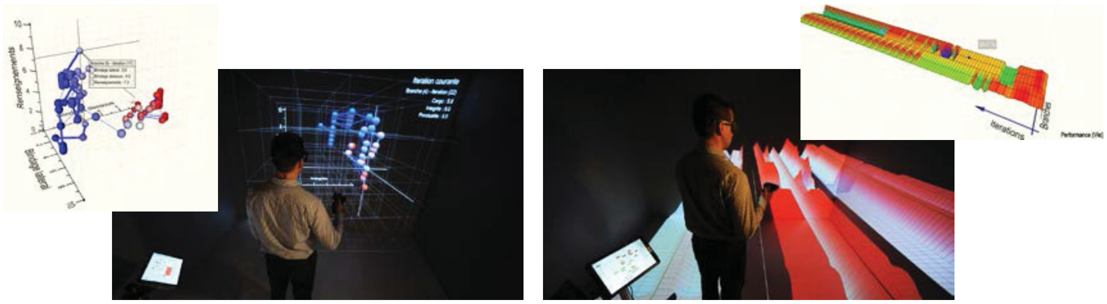


Figure 39: I-EXP V1 - Immersive tools at the SOP level (adaptation of desktop tools).

7.3 I-EXP V2

As aforementioned, the development of the I-EXP V1 toolset has been constrained by the cognitive experiments. IMAGE V2 is free of all cognitive experiment constraints, hence to the possibility to exploit VR concepts to their full potential within the I-EXP module and to immerse the user in a dynamic working VE he can build, organize and manage at will by exploiting a toolset for simulation, visualization and analysis. In this version, the I-SIM and I-EXP modules (at least the GUI) are combined within the same environment. More details can be found in (Mokhtari, Boivin, & Laurendeau, 2013; Mokhtari & Boivin, 2013b; Mokhtari et al., 2011).

7.3.1 Architecture

The I-EXP V2 engine is coded in C++ and is based on OSG. The I-EXP V2 environment can be used on a flat display wall or in a CAVE, both providing active stereoscopic visualization. The architecture also extends to desktop and multi-touch platforms, and supports collaborative work between users. The architecture has been designed to be independent of graphics and physics rendering engines, and is built as a tree, in which each branch is an entity, some entities having no correspondence with graphics and physics rendering trees. Three categories of entities have been defined:

- Manipulator: represents a hardware device (e.g. data glove, wand, keyboard / mouse…) used by human to interact with the objects in the environment. Several Manipulators of the same type can coexist in the I-EXP V2 environment;
- User: corresponds to a real user in the environment. It establishes the correspondence between the user and his Manipulator(s) and also manages object manipulation to avoid conflict between users;
- Object: represents an element that is not a Manipulator and with which User can interact. 3D Objects are special instances of Object: their pose (position, rotation, pivoting point and scale factor) can be modified and manipulations (move, rotate, zoom) can be applied on them.

Interaction is considered as a communication mechanism (i.e. flow of messages) allowing Manipulators to send interactions to Objects. Interactions are not directly coupled with Manipulators so the Manipulator design is independent of the Object design. Drag and drop is a mechanism that is implemented at the entity level (1) to transfer information from one Object to another (via a ghost concept) and (2) to instantiate associations between two Objects. Another flexible and powerful feature of the architecture is the MetaData concept which encapsulates all data types provided to the I-EXP module by the I-SIM module. The technical memorandum *IMAGE Exploration Module - Architecture* to be published in 2013 will report a more detailed description of the implemented architecture.

7.3.2 Human-centric immersive virtual environment prototype

The I-EXP V2 environment (see Figure 40 for an example) is entirely built, managed and controlled by the user during his work session (including object creation, manipulation, destruction, association of objects, etc.). The various objects available, which can be considered as a set of work tools, are presented succinctly in the following subsections.

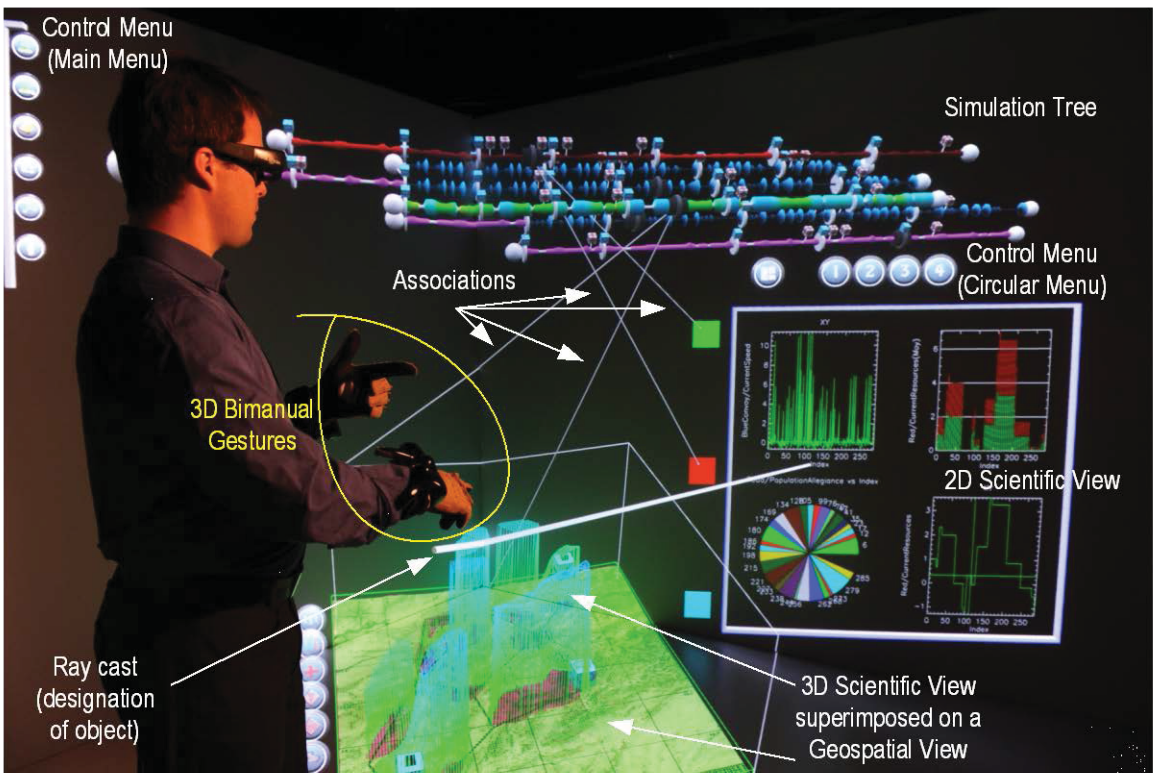


Figure 40: An instance of the I-EXP V2 CAVE environment.

Simulation Tree – One scenario, one parameter space

The Simulation Tree is the cornerstone of the I-EXP V2 environment. It is the direct link between the I-SIM and I-EXP modules. The Simulation Tree corresponds to a 3D implementation of the multichronic tree (visible on Figure 30), which is the visual representation of the simulation conceptual framework used by IMAGE V1. A first (basic) version of the Simulation Tree has been implemented in (Ouellet, 2012). The Simulation Tree encapsulates both (1) simulations logs in a dynamic layout and (2) simulation data. It can be related to the concept of hierarchy visualization (a branch of *Information Visualization*) whose visualization and layout techniques have been identified and listed in (Schulz, 2011)[5]. Within this classification, the Simulation Tree belongs to the category of 3D dimensionality, an explicit edge (node-link) representation and as exhibiting principal axis-parallel node alignment. Several Simulation Trees can coexist in the I-EXP V2 environment, each one representing one simulation space created by the user. To mark a point of interest on the Simulation Tree, the user can create sliders, objects that stick to the Simulation Tree and slide along its axes. Combining concepts of data and / or information visualization, the Simulation Tree provides users with (1) a visually explicit history of their experiments; (2) a self-explanatory representation of the simulations; and (3) interaction metaphors with running simulations for which parameters and resulting data, rendered in an adequate way, can assist users to better understand the simulated CS. Examples of Simulation Trees are shown in Figure 40 and Figure 41.

Geospatial View – A means of expliciting a simulation

The Geospatial View (visible in Figure 40 and Figure 41) allows the simulation status to be visualized as a snapshot at each simulation step (i.e. defined by a slider in the Simulation Tree). The Geospatial View corresponds to a terrain representation on which the scenario elements are displayed symbolically (cubing of the military symbology[6]). Information layers and gauges / indicators (being updated at each simulation step) can be added to inform the user of the status of simulation parameters. This tool also permits a simulation to be *played* (like a video). This object is active if and only if it is associated with a given simulation (in the Simulation Tree) by a drag-and-drop operation. The Geospatial View is conceptually similar to the simulation animation of the I-EXP V1 tool at the Tactical level (see section 7.2.2).

Scientific Views – Comparing simulations in different ways

At any time, to visualize, analyze and explore simulation data or compare simulations, the user can configure Scientific Views (already associated or not to one or several simulations) according to different plot functions such as: (1) 2D plots - line (plot), bar (histogram, stacked), and area (pie) graphs; the user can create up to 4 different graphics within a figure window, each representing different data combinations; and (2) 3D plots: similar to the 2D plots but in 3D.

To associate a simulation to one Scientific View, users can drag and drop this simulation onto a Scientific View and, by a set of menus, select data they wish to visualize and / or analyze (and compare) as well as the type of plot (2D or 3D) to be used. Scientific Views are based on the PLplot library (http://plplot.sourceforge.net/). Examples of Scientific Views can be seen in Figure 40 and Figure 41.

Association – Communication channel between objects

In order to create an environment in which objects can interact with each other, the concept of *association* has been implemented at the entity level. This mechanism allows two objects to

[5] For additional detail, the reader can also consult the Tree Visualization Survey webpage: treevis.net.
[6] Symbols and colors are in conformity with the MIL-STD-2525C Common Warfighting Symbology (US DoD) and with the NATO Standard APP-6(B).

communicate and exchange data in an asynchronous way using a common protocol. In the environment, an association is represented graphically as a physical link between objects (a bidirectional arrow). Objects' associations can be deleted by the user when they are no longer needed. Associations are shown in Figure 40 and Figure 41.

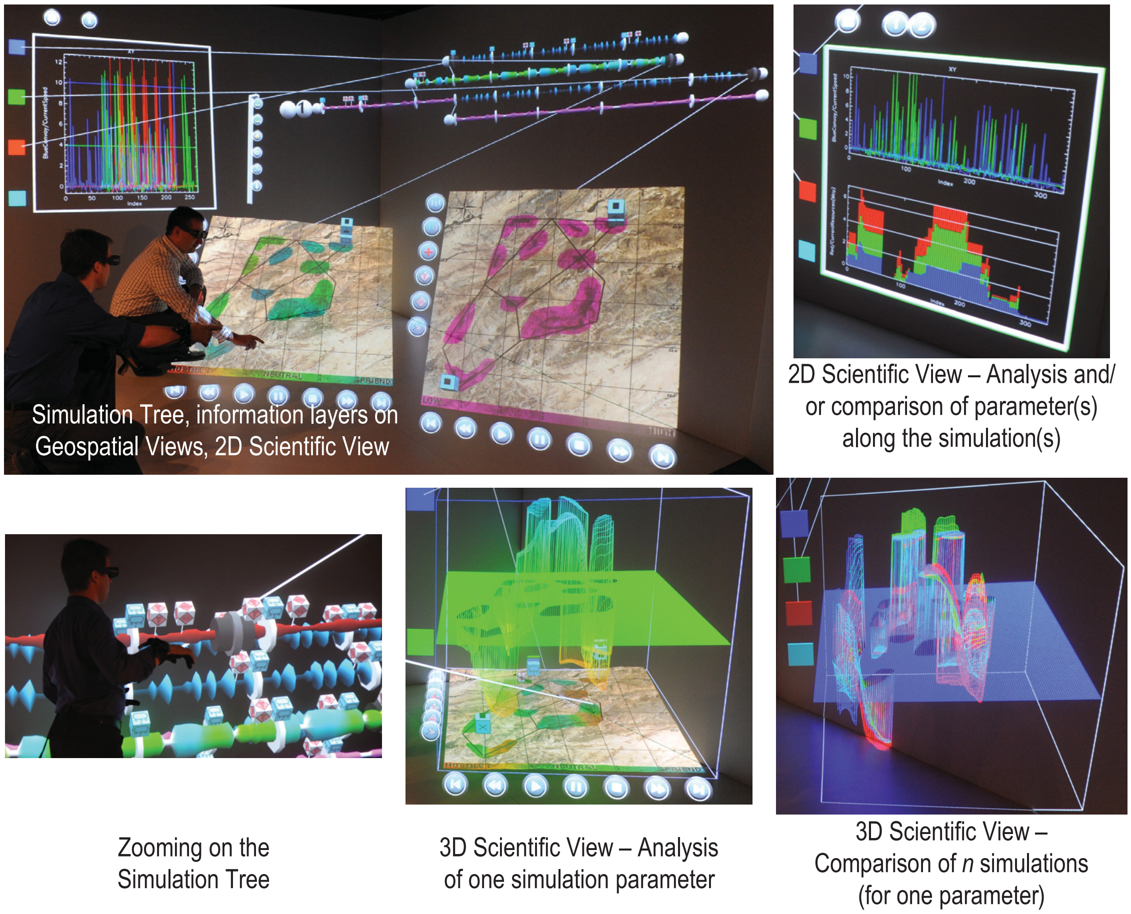


Figure 41: Work tools available in the I-EXP V2 CAVE environment.

Control menus

Two types of control menus (see Figure 42) are available to users:

- Main menu: is a traditional linear menu composed of a set of icons that represent actions that can be either selected or dragged onto the target object; is always visible; allows to connect the I-EXP module to the simulator, to create objects (some by changing the value of parameters), and to associate information with elements of the environment. Note: object deletion is achieved by throwing the object above the head of the user (the 3D bimanual gestural interface for achieving this operation – and others – is explained in section 7.3.3).
- Circular menu: is a hierarchical menu that can be attached to specific objects for additional configuration capabilities (Huot & Lecolinet, 2007); is visible on-demand,

and allows choosing which and how data should be visualized and displayed. A circular menu is an interesting technique for displaying many items in a confined space. Besides, it speeds up and optimizes the display options without causing information overload. The item selection time is constant due to the circular organization around the activation point unlike linear menu where the item selection time varies depending on the item position in the menu (Bailly, 2009).

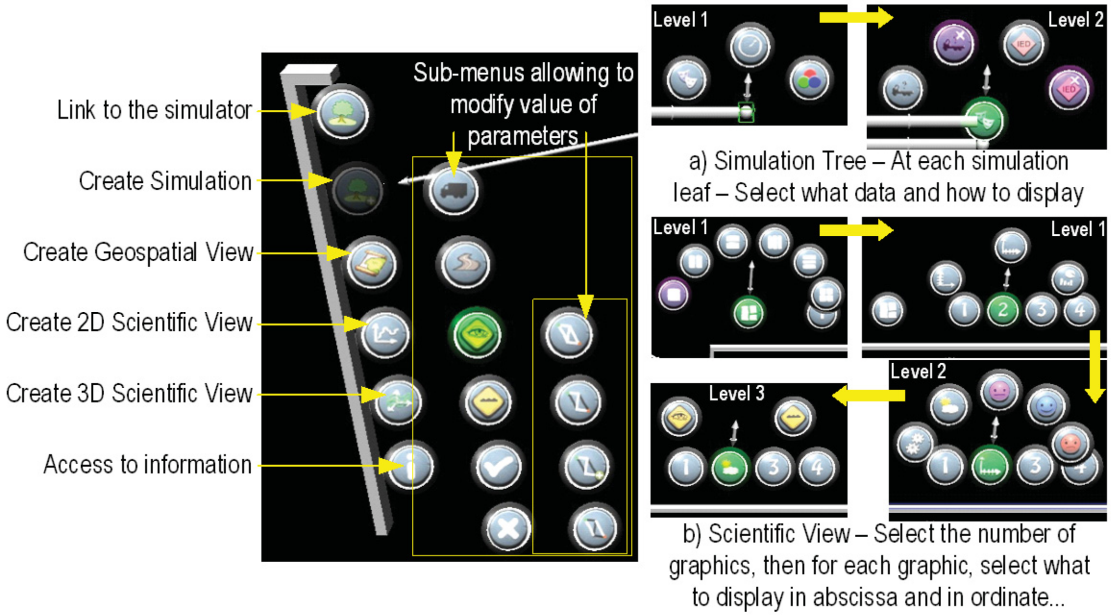


Figure 42: Control menus of the I-EXP V2 CAVE tool - (left) main menu (+submenus) and (right) examples of circular menus (+part of their hierarchy).

7.3.3 3D bimanual gestural interface

Gestural interfaces have attracted considerable attention in VR from the very beginning, since hands are used naturally by humans to interact with their immediate environment. Consequently, using hands as a means of controlling an IVE is often considered in VR. Bimanual interfaces, either gestural or not, have also received a significant amount of attention in recent years. It is now acknowledged that, with carefully designed interactions, two hands can perform better than one on a given task (Hinckley, Pausch, Proffitt, & Kassell, 1998; Owen, Kurtenbach, Fitzmaurice, Baudel, & Buxton, 2005; Veit, Capobianco, & Bechmann, 2008).

To enable the interaction between the user and the environment, a 3D bimanual gestural interface using Cybergloves has been developed. It is built upon past contributions on gestural interfaces and bimanual interactions to create an efficient and intuitive gestural interface tailored to IMAGE needs. More details can be found in (Lévesque, Laurendeau, & Mokhtari, 2011, 2013; Lévesque, 2011).

Related work

A major contribution in the domain of bimanual interactions is the "kinematic chain model" proposed by Guiard (1987). By studying examples of real-world human bimanual interactions (e.g. writing on a sheet of paper, opening a bottle), Guiard observed that work is split between the two hands in a structured way. He established three principles that describe asymmetric bimanual interactions (in the case of right-handed users): (1) the right hand operates in a spatial frame of reference that is relative to the left hand, (2) the left and right hands operate on different spatial and temporal scales of motion and (3) the left hand usually precedes the right hand in action. Hinckley et al. (1998) have developed a bimanual interface for 3D visualization that uses Guiard's right-to-left reference principle with great efficiency. Veit et al. (2008) later validated these theories in a 3D environment. On the topic of gestures, Nielsen, Störring, Moeslund, and Granum (2004) have extensively covered the design and selection of gestures for an interface, providing ergonomical insight as to which type of gestures should be used in gestural interfaces. In their bare-hands 3D user interface, Schlattmann and Klein (2009) used the distance between hands for selecting and manipulating objects.

Gestural interface design

The interface is based on Guiard's "kinematic chain model", not in the way that objects are manipulated, but rather in the way interactions are initiated and completed by the hands. The following guidelines were used in the design of the interface: Guiard's principles for bimanual interactions are respected; the right hand executes the interactions demanding greater precision; the number of gestures to be used is kept to a minimum since there is a learning curve imposed on users for remembering all gestures; whenever possible, actions should be mapped to gestures that are semantically related. It is also important to give users proper instructions (cues as to why a given gesture was chosen) and training; the hand gestures used in the interface reuse some ergonomic guidelines found in (Nielsen, Störring, Moeslund, & Granum, 2004), e.g. avoiding outer positions and relaxing muscles as much as possible; the beginning of interactions is associated with muscle tension and the end with the release of the tension (as proposed in (Cerney, 2005)); and finally the interface is designed for expert users.

This interface uses *static* gestures, i.e. static hand configurations or hand postures (e.g., a clenched fist), not to be confused with *dynamic* gestures (e.g., a waving hand). The left hand plays the role of a mode switcher while most of direct manipulations are executed by the right hand. Hand gestures are executed by the left hand to select interaction modes, while the right hand performs the interaction itself, e.g. pointing, moving or scaling the desired object. The selected gestures used in the I-EXP V2 CAVE environment are presented in Figure 43.

A support vector machine (SVM) is responsible for gesture recognition. SVMs provide a performance which is similar both in execution time and accuracy (Heumer, Amor, Weber, & Jung, 2007) to that of neural networks and do not require a network topology to be determined. The SVM can recognize user's gestures with good accuracy and precision for sets composed of around 10 gestures from a pre-established set of gestures with which it was trained beforehand.

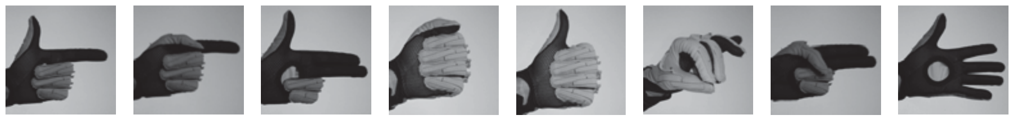
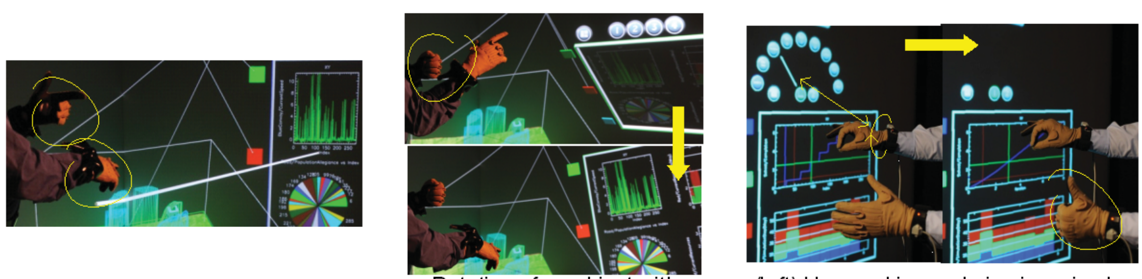


Figure 43: Hand gestures used in the I-EXP V2 CAVE environment.

3D interactions

User actions in the environment have been separated into 4 categories of gestures: (1) designation (based on ray-casting) and selection of objects; (2) generic manipulations, which group all interactions related to moving / positioning, resizing and rotating objects; (3) specific manipulations, which are interactions tuned for specific objects (e.g. the playback capabilities associated to the Geospatial View – allowing for control of the playback using the right hand's gestures and or pose while the left hand maintains the mode); and (4) system control, which represents all actions that are related to (main and circular) menus and modifying the way the environment behaves. Symbolic input of numerical values is also implemented through a behaviour similar to that of circular menus, i.e., by rotating the right hand and accepting or cancelling with the left hand. Action examples are presented in Figure 44. In the latter case, circular menus expand once selected and the right hand's orientation can then be used to select a menu option. The user confirms or cancels a selection with his left hand by pointing his thumb either up or down. To limit fatigue issues, the thumb-down gesture is not required to point all the way down but only half-way.

There is no need for travelling interactions in IMAGE V2 since the user is located at the center of the IVE and has access to objects without *really* navigating in the environment (the user can move freely in the CAVE environment which is a restricted area and consequently his motion cannot be considered as a travelling interaction).



Designation of an object with the left hand and the right hand pointing

Rotation of an object with the left clenched fist and the right hand pointing

(left) User making a choice in a circular menu – *pinch,* and (right) user confirming this choice – *thumb-up*

Figure 44: Examples of (left) selection action; (middle) generic manipulation action; (right) system control action.

7.3.4 Multi-touch interaction metaphor

In adapting I-EXP V2 to a multi-touch screen, three specific aspects have been addressed: (1) to ensure both collaborative (accommodate multiple users in the same working environment) and individual use; (2) to experiment and test the flexibility of the I-EXP V2 architecture, i.e. its adaptability to a spectrum of devices (i.e. desktop, multi-touch displays, and IVE such as CAVE) in terms of visual and interaction aspects; and (3) to exploit natural user interaction (ease of use) provided by multi-touch devices.

Multi-touch technology has attracted a lot of attention in recent years mainly by its potential to improve human computer interaction. Multi-touch refers to a touch sensing surface ability to recognize the presence of two or more points of contact with the surface. Multi-touch sensing devices are inherently able to accommodate multiple users simultaneously (especially useful for larger interaction applications such as interactive display walls and tabletops) (Buxton, 2007)[7]. The platform used is a 3M M2256PW 22in 1680×1050 20 finger multi-touch Widescreen LCD Monitor (http://solutions.3m.com/, section *Touch systems and Touch screens*).

The gesture language developed to communicate with the workspace environment is summarized in Figure 45:

- Gestures using one contact: (1) select an object (*Tap*); (2) hold for .5 second an object enables "drag and drop"; (3) move an object from point A to point B on the multi-touch screen plane (*Pan*); and (4) change the numerical value of parameters when creating a simulation (*Scroll*).
- Gestures using two contacts: (1) resize with two fingers (*Pinch* and / or Two Hand *Pinch – Spread* and / or Two Hand *Spread*); (2) rotate an object around an axis normal to the multi-touch screen plane (*Rotate* and / or Two Hand *Rotate*).
- Gesture using three contacts: rotate an object around an axis parallel to the multi-touch screen plane (One Hand *Rotate*).
- Gesture using four contacts: holding an object for two seconds forces its destruction.

When several fingers touch the multi-touch screen, they do not belong necessarily to the same hand. To succeed "drag and drop" and destruction actions, the fingers should not move after contact. Examples of actions are shown on Figure 46.

[7] Buxton's multi-touch Web page provides a thorough overview of the underlying technologies and history of multi-touch surfaces and interaction.

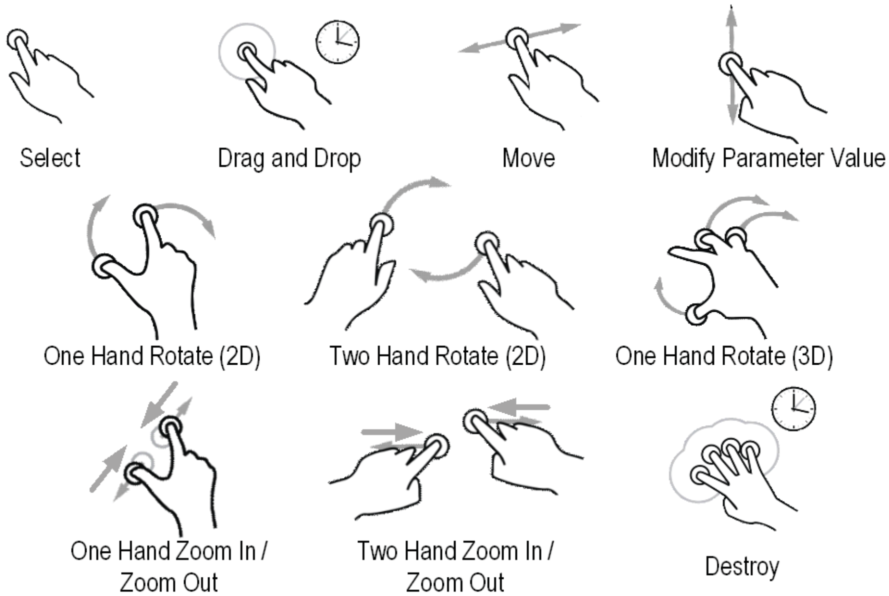


Figure 45: Gesture language developed for the I-EXP V2 multi-touch screen interface.

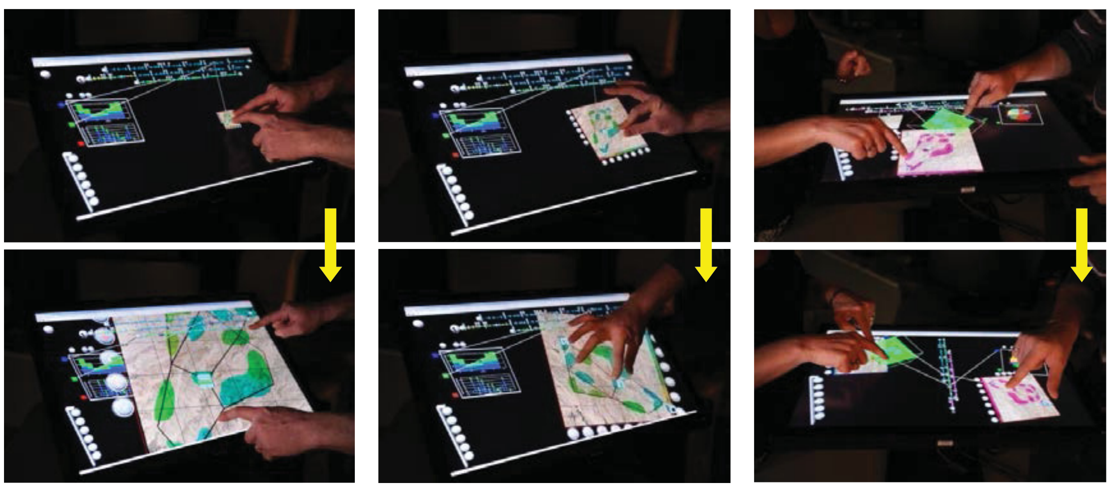


Figure 46: Examples of actions in the I-EXP V2 multi-touch tool : (left) resize an object – two hand rotate, (middle) rotate and resize an object – one hand rotate and zoom-in, (right) exchange of objects between two users while the display is configured as a table-top.

7.4 Discussion

The tools developed for the I-EXP module are prototype tools and therefore improvements are needed to make them more efficient to users. In this spirit, the I-EXP V1 Tactical tool (section 7.2.2) could be improved to include visual and analytical comparison for a set of simulations (in a way different from Tableau). The access to the I-EXP V1 Tactical tool by the I-EXP V1 SOP tools should be made possible in order to make available the data hierarchy to users. Some improvements could be brought to the synergy between the I-SIM and I-EXP modules in order to propose an integrated and more ergonomic tool. In addition, the hybrid toolset (including immersive and non-immersive modes, section 7.2.3) could be designed differently to enhance the switch between the two modes and a better fluidity in the tool for users. The I-EXP V2 (section 7.3.2) is a proof-of-concept and consequently there is room for improvement. For example, the Simulation Tree could be augmented to include flexibility / self-configurability to answer requests from users. The team is currently exploring ways of connecting several CAVEs in order to implement collaborative work in IVEs in the context of IMAGE and others projects. In the proposed 3D bimanual gestural interface (section 7.3.3), the fact that the left hand defines the mode of action means it will always act before the right hand, thus respecting the precedence principle of Guiard. The left hand also acts on different spatial and temporal scales than the right hand (typically, once at the beginning and at the end of every interaction), responding to Guiard's second principle as well. The principle of right-to-left spatial reference should only be met if there were an interaction that is complex enough to justify its need, which is not the case yet for IMAGE. However, in a way, the right hand depends on the left hand to define its mode of operation so it could be claimed that the right-to-left reference principle is also met, although this is still open for discussion because it is not a spatial reference per se. In the proposed interface, hand gestures are captured by data gloves, rather than computer vision, for increased efficiency and reliability, but future work could study the use of recent technology advancements in bare-hand gestural interfaces like the Microsoft Kinect. Finally, a data exchange pipeline could be built between the I-EXP and I-REP modules.

This page intentionally left blank.

8 Experimental validation

The study was designed to assess how the human comprehension capacities of complex military situations would be augmented with IMAGE suite of tools in comparison with usual tools. More precisely, the research questions were:

- Does IMAGE improve comprehension of CS?
- What is the added value of immersion for IMAGE?
- What are the typical and the efficient ways to exploit IMAGE?

This chapter presents a summary of the method and results. A report on IMAGE experiment (Gagnon et al., 2012) contains more details on many aspects presented here. Nevertheless, the current chapter provides additional results and interpretations that are not included in the experiment report.

8.1 Method

Three experimental platform conditions were tested: Baseline, IMAGE-Desktop and IMAGE-CAVE. Each of these experimental platforms came with a set of common but also different software applications and hardware. Participants in all conditions studied the CS described in Chapter 3.

Military analysts were judged as the ideal candidate for this experiment, but because of the lack of availability of such experts, civilian participants were selected in order to get statistical significance. The experiment was conducted mainly at Université Laval because of its proximity with participants but also at DRDC Valcartier for IMAGE-CAVE condition where the CAVE was located. The sample was composed of 58 persons, either graduate or undergraduate students, with four semesters completed in the domain of operational research, computer science, and other related fields. The study has been approved by the DRDC Ethic Committee and the Ethics committee of Université Laval. All individual participants in this study gave written informed consent prior to their participation.

Software

Participants in IMAGE conditions were provided with the IMAGE V1 modules described in chapters 4 to 7. Baseline condition participants were using standard commercial tools that they would normally use for this kind of analysis. For the Simulation tools, a simple version of Multichronia was provided in Baseline; the simulation capability was limited to one branch at a time. Microsoft Excel was the main Exploration tool for Baseline condition and was completed with a Java application showing an orthographic top-view of the convoy at the Tactical level. Tableau and Eye-Sys were the Exploration tools for the IMAGE conditions. Microsoft Word and IMAGE-CoGUI (i.e. I-REP V1) were the Representation tools used in Baseline and IMAGE conditions respectively. Finally, a Question Manager application was included in all conditions to display instructions, remaining time, and exploration questions, as well as to collect participants' answers (see Annex A). Table 2 shows the deployment of tools on each screen for the three conditions.

Table 2: Tools deployed on each screen according to the experimental conditions.

	Left screen	Middle screen	Right screen	Tablet PC
Baseline	Word	Multichronia Light Question manager	Eye-Sys Light Excel	
IMAGE-Desktop	CoGUI Word	Multichronia Question manager	Eye-Sys Tableau	
IMAGE-CAVE	Question manager	Eye-Sys Tableau	CoGUI Word	Multichronia Eye-Sys*

* The tablet PC allowed changing views and manipulation of data in immersive mode.

Hardware

Two hardware set-ups were used for the experiment (Figure 47). Baseline and IMAGE-Desktop shared the same computer platform: a BOXX 7400 workstation (2x Opteron 275 @ 2.2 GHz, 4 GB RAM, 250 GB HD, 2x NVIDIA Quadro FX 4500, Windows XP) with three 22-inch flat screens with a resolution of 1600×1200. IMAGE-CAVE used two BOXX 7500 (2x Opteron 2222 @ 3 GHz, 4 GB RAM, 146 GB HD, 2x NVIDIA Quadro FX 5600G, Windows XP) and a Safari Tablet PC connected to a CAVE. The CAVE consisted of a three-wall one-floor stereoscopic immersion environment called the Flex (http://www.mechdyne.com). Four Digital Projection (http://www.digitalprojection.com) Highlite 8000Dsx+ digital light processing projectors displayed images at 8000 ANSI lumens and a native resolution of 1400x1050@96 Hz on the 10'8"×8' screens. Participants wore RealD Crystal Eyes 3 (http://www.reald.com) shuttle glasses. An Intersense IS-900 VET motion tracking system (http://www.intersense.com) tracked the Wand device and the participant's position for correcting the image perspective when in immersion mode.

Figure 47: Experimental set-up: Baseline and IMAGE-Desktop (left), IMAGE-CAVE non-immersive (middle) and IMAGE-CAVE immersive (right).

Measures

Measuring comprehension is challenging. As there is equipment to measure blood pressure or brain activities, there is no scientific apparatus to quantify how much a person understands, or more precisely perceive, the reality of a situation. Therefore, there is a need to indirectly estimate comprehension, measuring something else, a proxy, which represents well a level of understanding. For instance, asking people to explain in their own words what they understood is a way but is tedious to analyse. Techniques such as the one used by (O'Connor et al., 2004) is easier to analyse and pretty interesting. They conducted a task analysis to identify concepts used by people to describe a situation. Then, they asked experts to use these concepts for constructing a map representing their individual understanding. Finally, expert concept maps were compared to identify commonalities, or sharedness to be used by participants in creating their individually constructed mental model. The limitation of this approach is that it relies a lot on the communication capability of the participants. But moreover, it would have been too close to one of the tool under investigation, I-REP, and could bias the results. Performance was selected as the most appropriate for the current study used complemented by two other types of measure. Three measures of comprehension and one measure of error were developed to quantify the gain in comprehension provided by IMAGE:

- Analysis performance score that evaluated the correctness of answers to four questions at the Tactical level and three questions at the SOP level while participants have access to their tools. Stated otherwise, it is a measure of how well participants could do their job with the tools. The score is the average mark to each question. Questions focus on complex aspects of the CS. For the SOP level, scores are normalized between 0 (worst) and 1 (best) with a non-linear transformation that compensates for the frequency of occurrence (i.e. high values of MoPs are far less frequent and, since a small increase of a high value is more difficult, a correction considers this aspect).
- Prediction score that evaluates answers to ten two-choice questions based on the knowledge participants had acquired during the analysis phase. Scores are normalized between 0 (worst) and 1 (best).
- Quality of mental model score that evaluates the long-term associative knowledge (correlation) between the most important variables of the CS. This measure captures the participants' long term associative knowledge. Participants had to provide coefficient between -1 and 1 for each pair of variables. The Quality of mental model score was based on two intermediate measures. First, a close-to-ideal measure was calculated by averaging the square distance between participant's response and the real (calculated) correlation matrix. The second measure repeated the same process but this time by comparing the participant's matrix with a matrix as if participant had answered randomly. This "random" matrix had only null values since, on the average, the mean value between -1 and 1 is zero. The result was called the far-from-random measure. The Quality of mental model score was then obtained by dividing the close-to-ideal by the far-from-random measures. This score represents how good a participant was at estimating correlations compared to obtaining these values purely by chance.
- Response selection fallacy (RSF) rate that offers the opportunity to evaluate the "error-reduction" potential of a suite of tools. One of the tasks of the participants is to explore the parameters space by instantiating simulations and then pick one of them as an answer to analysis questions. At the SOP level, participants do not necessarily choose the best instantiated simulation to answer the three questions related to finding the combination of parameters that maximizes each MoP. Similarly, at the Tactical level, participants can provide answers that are wrong while the right answer is in the pool of simulation they create or they provide less optimal solutions than those in the pool. The frequency of these errors, the RSF rate, is calculated as the proportion of simulations instantiated by a participant that better answer the questions than the one those he provided.

Additional but less important measures like user interaction metrics, tool functionalities usage, and analysis strategy were also included and will be explained in due course in the following sections.

Procedures

The 58 participants were randomly assigned to one of the three conditions: Baseline (15), IMAGE-Desktop (22) and IMAGE-CAVE (21). Participants underwent between 9 hours (Baseline condition) and 14 hours (IMAGE conditions) of experiment distributed over six or seven sessions (see Table 3). During their first session, participants familiarized themselves with the tools. Afterward, participants received a theoretical course on complexity, a description of the convoy CS (see Chapter 3), and notions on complexity and operational research analyst tasks. Then, participants in IMAGE conditions devoted 45 minutes to develop an initial comprehension model of the CS using I-REP V1. After, participants completed a cycle of two sessions of investigations dedicated to studying the CS Tactical and the SOP levels. Participants had to play the role of an operational research analyst (Richardson, Mathieson, & Cilliers, 2000). The same investigation process was repeated in a second cycle to observe the effect of learning. Each session comprised a 75-minute analysis phase during which participants used the tools to answer questions, a 15-minute prediction phase for answering questions without having access to the tools, and, for IMAGE conditions, an additional 30-minute comprehension model update. Afterward, participants in IMAGE conditions underwent a think-aloud protocol session and filled usability questionnaires. Finally, few weeks after the last session, participants passed an on-line associative knowledge test for measuring the quality of their mental model. On the total difference in time (5 hours) between the conditions, 3 hours related to either tutorial activities or post measures were not detrimental to Baseline condition. As for the additional 2 hours dedicated to representation, its benefits were confounded with those provided by the improved tools. The experiment aimed at measuring not only the benefit of the tools but also those of IMAGE process for which no equivalents were found in Baseline for some tasks (like representation). In summary, the current experiment design was the best and fairest compromise for the given experimental objectives.

Table 3: Experimental procedure.

			Cycle 1		Cycle 2		
			Tactical	SOP	Tactical	SOP	
Baseline	Tutorial Scenario *1 hour*		Analysis Prediction *2 hours*	Analysis Prediction *2 hours*	Analysis Prediction *2 hours*	Analysis Prediction *2 hours*	
IMAGE	Tutorial	Tutorial Scenario Representation	Analysis Prediction Representation	Analysis Prediction Representation	Analysis Prediction Representation	Analysis Prediction Representation	Think-aloud Questionnaire
	1 hour	*1.5 hour*	*2.5 hours*	*2.5 hours*	*2.5 hours*	*2.5 hours*	*1.5 hour*

8.2 Performance results

The improvement in comprehension based on performance score were analysed by combining IMAGE-Desktop with IMAGE-CAVE, so that the first and most important research question could be answered with greater statistical power.

8.2.1 Analysis performance score

The score at each question was averaged to create the analysis performance score, which was averaged for each condition, level, and cycle (see Table 4).

Table 4: Mean (standard error) analysis performance core by cycle, level, and condition.

	Cycle 1		Cycle 2		AVG (cycle)		AVG
	Tactical	SOP	Tactical	SOP	Tactical	SOP	overall
IMAGE	.51 (.028)	.77 (.013)	.59 (.024)	.84 (.011)	.55 (.022)	.81 (.011)	.68 (.014)
Baseline	.42 (.069)	.72 (.027)	.55 (.067)	.79 (.025)	.48 (.024)	.76 (.042)	.62 (.030)

Results showed a systematic increase in mean analysis performance across cycles both for the Tactical $F(1,50) = 14.421$, $p < .001$ and the SOP $F(1,50) = 33.698$, $p < .001$ levels. The analyses revealed no interaction between cycles and experimental conditions for both Tactical $F(2,50) < 1$, N.S. and SOP $F(2,50) = 1.161$, N.S. levels. Results reveal a significant higher analysis performance score for IMAGE ($M = .68$, $SD = .09$) over Baseline ($M = .62$, $SD = .11$) conditions; $t(51) = 2.076$, $p = .022$. The average score increased between the first and the second as expected. Overall, IMAGE contributes to improve on-line analysis but the effects at specific levels of complexity were not statistically significant due to large variability between participants. Nevertheless, the difference in score observed between Baseline and IMAGE remained when the level of complexity was increased and after a second cycle as illustrated in Figure 48. Results do not support the acceleration of comprehension by IMAGE but data suggest that such acceleration is more likely to be found in less CSs.

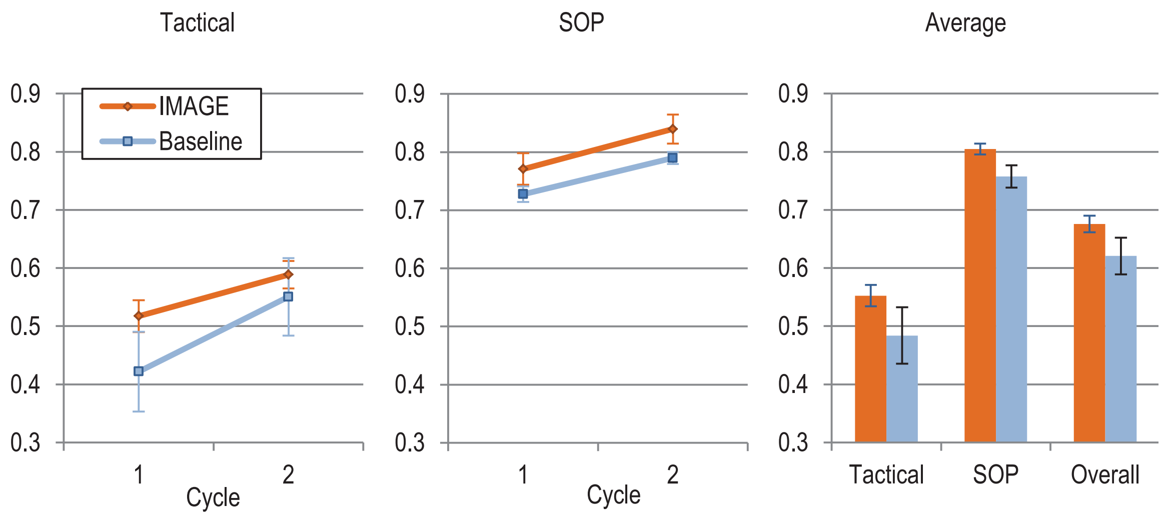


Figure 48: Mean analysis performance score by cycle, level, and condition (error bars show standard errors).

8.2.2 Prediction performance score

After the 75-minute analysis session, participants were asked a series of 10 predictive questions without having access to IMAGE or Baseline tools. Average scores are presented in Table 5 and Figure 49.

Table 5: Mean (standard error) prediction performance core by cycle, level, and condition.

	Cycle 1		Cycle 2		AVG (cycle)		AVG
	Tactical	SOP	Tactical	SOP	Tactical	SOP	overall
IMAGE	.64 (.024)	.56 (.025)	.66 (.018)	.53 (.021)	.65 (.015)	.54 (.016)	.60 (.016)
Baseline	.63 (.025)	.63 (.030)	.48 (.031)	.58 (.043)	.56 (.024)	.62 (.026)	.58 (.011)

A two-way analysis of variance for each level (Tactical and SOP) was conducted using a Bonferroni adjusted alpha levels of .025. The analyses yield a main effect for condition at the Tactical level, $F(1,54) = 10.4$, $p = .002$ but no main effect at the SOP level, $F(1,51) = 3.33$, N.S., indicating that the prediction score was significantly higher for IMAGE ($M = .65$, $SD = .014$) than Baseline ($M = .56$, $SD = .025$) condition but only at the Tactical level. The same analysis revealed a main effect for cycle at the Tactical $F(1,54) = 6.27$, $p = .015$ but not at the SOP $F(1,51) = 2.03$, N.S. levels. Finally, analyses revealed an interaction between condition and cycle at the Tactical $F(1,54) = 10.12$, $p = .002$ but not at the SOP $F(1,51) = .136$, N.S. level.

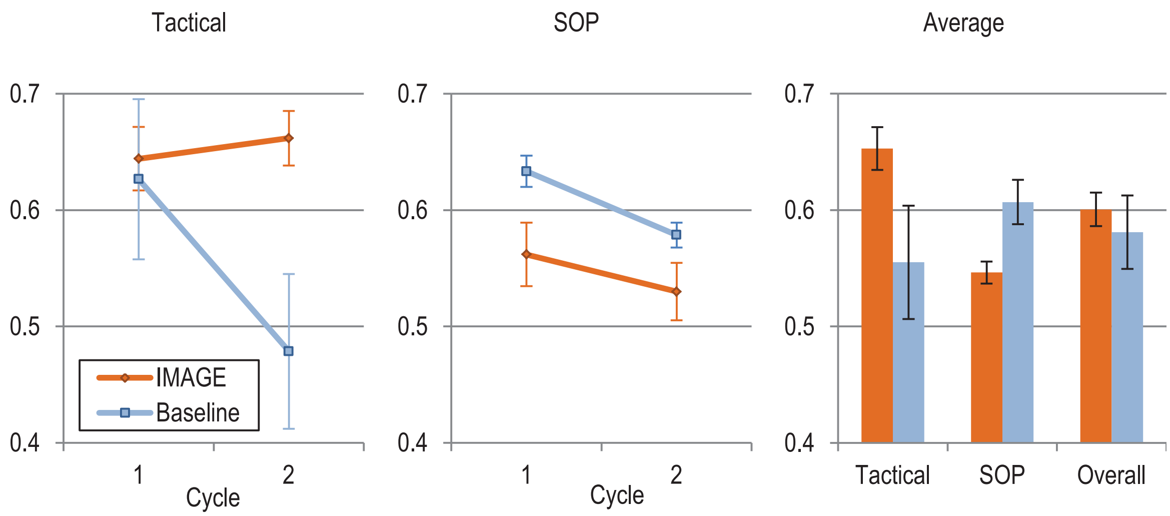


Figure 49: Mean prediction performance score by cycle, level, and condition (error bars show standard errors).

The results for the prediction performance are not showing the benefit of IMAGE as clearly as the results of the analysis performance score. Overall, there was no difference between the conditions, $t(56) < 1$, $p = .35$. Improving comprehension when no tools can be consulted, i.e. based on knowledge previously acquired, does not seem warranted; IMAGE provides some prediction benefits for less CSs but still fail to provide improvement for more complex ones.

8.2.3 Mental models

The Quality of mental model was based on the Pearson product-moment correlation coefficient calculated for each combination of the eight variables at the Tactical level and seven variables at the SOP level. Figure 50 shows correlation matrix calculated from the 726 simulations covering 100% of the parameter space at the Tactical level and from 750 000 simulations sampling strategically the parameter space by using the scrambled Halton sequence approach (Bhat, 2003) for the SOP level.

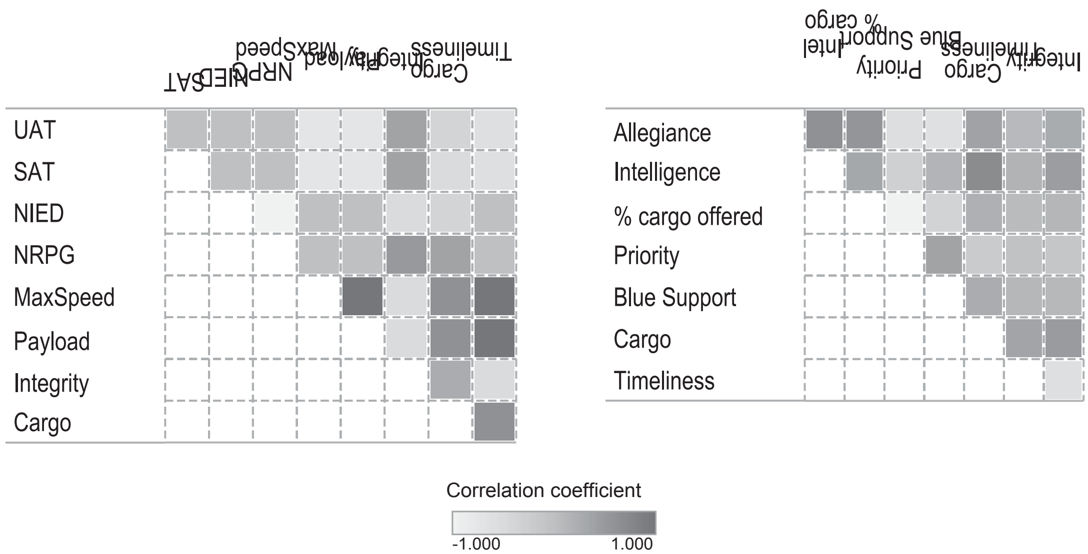


Figure 50: Calculated correlation matrix for Tactical and SOP levels.

Figure 51 shows the average value of the Quality of mental model score for both conditions and cycles.

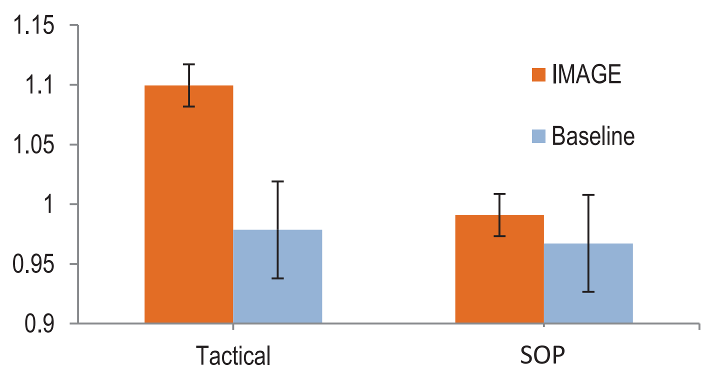


Figure 51: Quality of mental models scores by experimental conditions and complexity levels (error bars show standard errors).

At the lowest level of complexity (Tactical), participants in the IMAGE condition exhibited higher mean Quality of mental models scores than the participants from the Baseline condition $t(38) = 2.341$, $p = .013$. At the highest level of complexity (SOP), the observed difference in mean Quality of mental models score across IMAGE and Baseline condition was not statistically significant $t(37) < 1$, N.S.. Here again IMAGE provides a benefit but only at the lowest level of complexity.

8.2.4 Response selection fallacy

RSF rates were calculated for each condition, cycle, and level as shown in Table 6 and Figure 52.

Table 6: Average RSF rates (standard error shown in parenthesis).

	Cycle 1		Cycle 2		AVG (cycle)		AVG overall
Condition	Tactical	SOP	Tactical	SOP	Tactical	SOP	
IMAGE	.04 (.010)	.19 (.031)	.01 (.004)	.06 (.014)	.03 (.005)	.13 (.018)	.08 (.010)
Baseline	.11 (.036)	.25 (.059)	.08 (.035)	.12 (.036)	.10 (.025)	.19 (.036)	.14 (.023)

A two-way analysis of variance for each level (Tactical and SOP) was conducted using a Bonferroni adjusted alpha levels of .25. The analyses yield a main effect for condition at the Tactical level, $F(1,56) = 10.88$, $p = .002$ but not at the SOP level, $F(1,55) = 2.73$, N.S., indicating that the RSF rate was significantly lower for IMAGE ($M = .027$, $SD = .051$) than Baseline ($M = .096$, $SD = .137$) but only at the Tactical level. The same analysis showed a systematic decrease of RSF rate across the cycle for the Tactical $F(1,56) = 7.24$, $p = .009$ and the SOP $F(1,55) = 16.54$, $p < .001$ levels. No interaction was found between cycle and experimental conditions at the Tactical $F(1,56) = .11$, N.S. and at the SOP $F(1,55) = .001$, N.S. levels. Overall, a statistically significant difference on RSF rate was observed between IMAGE and Baseline conditions $t(56) = 3.1$, $p = .002$, such that the average RSF rate was significantly lower (half the value) for IMAGE ($M = .076$, $SD = .063$) than Baseline ($M = .142$, $SD = .090$) condition.

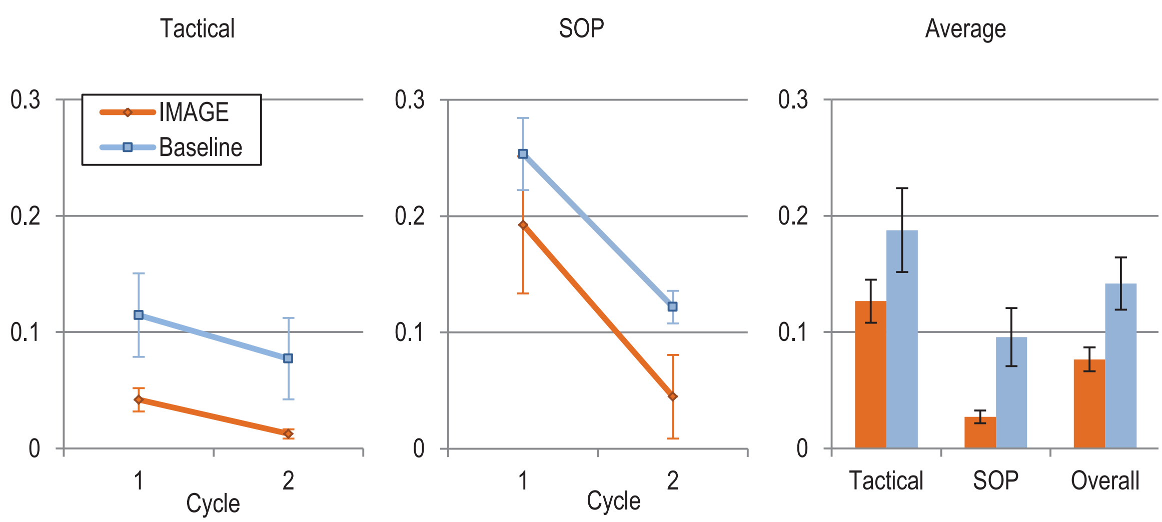


Figure 52: Averaged RSF rates by cycle, level, and condition (error bars show standard errors).

The difference in performance score (analysis phase) between Baseline and IMAGE is due to the additive effects of the RSF rate and the level of comprehension as shown in Eq. (1). If one of the two right-hand side terms could be found, the other would be deducted by subtracting it from the analysis performance score presented in previous section.

$$\Delta Perfo_{Total} = \Delta Perfo_{RSF} + \Delta\, Perfo_{Comprehension} \quad (1)$$

The fact that Baseline and IMAGE conditions have different average RSF rate complicates the task. If it can be demonstrated that the difference of performance between IMAGE and Baseline is the same for any hypothetical RSF rates, then the calculus would be simplified. The following method aimed at demonstrating such constant gap for the SOP level only since questions asked at the Tactical level (good / wrong) do not allow ranking simulations. First, various hypothetical RSF rates were emulated by selecting the n^{th} best simulation instantiated by a participant. Then, an analysis performance score was calculated for this simulation instance. For instance, if a participant had created 50 simulations, an analysis performance score for a RSF rate of .1 (or 10%) was calculated with the fifth best simulation among the 50. Figure 53 shows such an emulated guide analysis performance score averaged for all participants for a series of hypothetical RSF rates. The difference in score between both conditions (*M* = .022, *SD* = .003) is almost constant between .03 and .25, the region where are located the RSF rates of most participants. The vertical difference can be explained by an improvement of understanding since the RSF rates are artificially fixed equals.

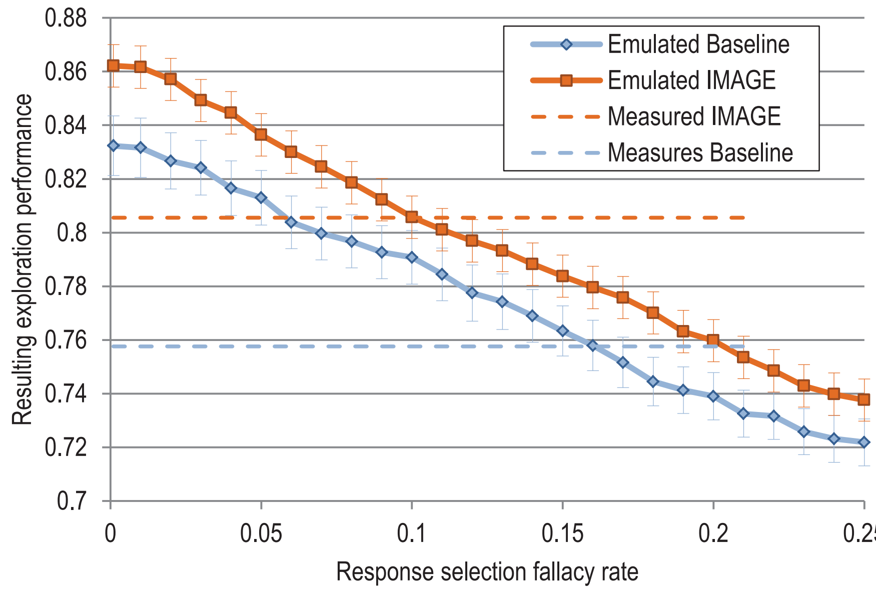


Figure 53: Calculated analysis performance score from speculated RSF rate for SOP level (error bars represent standard error).

Comprehension explains approximately .022 out of .048 (.806-.758) of the gain of guided analysis performance score provided by IMAGE at the more complex level, that is to say 45% of gain in performance provided by IMAGE comes from a better comprehension, the remaining 55% is provided by the better RSF-reducing potential of IMAGE (see Figure 54). These values are approximations since the current method averages values in a context where the relation between those values is not perfectly linear.

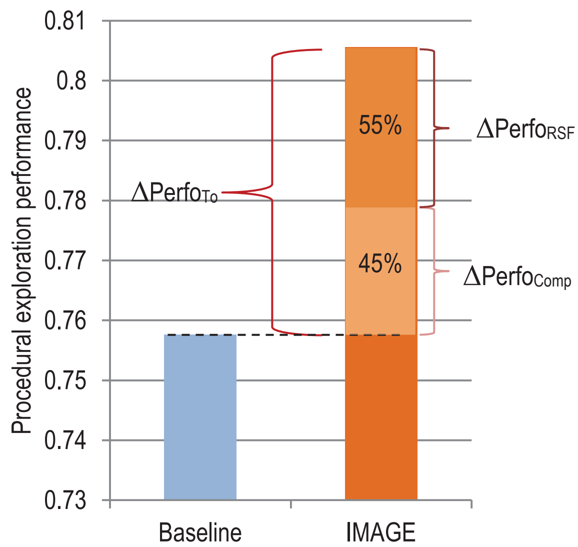


Figure 54: Contribution of RSF and comprehension to the difference of analysis performance score between IMAGE and Baseline conditions.

8.2.5 Immersion added-value

It was hypothesized that immersion would provide an additional benefit over the desktop version. Multi-dimensional visualization in an immersive context would allow participants to better understand relation between data as showed in previous studies (Arms, Cook, & Cruz-Neira, 1999; Raja, Bowman, Lucas, & North, 2004; Van Dam, Forsberg, Laidlaw, LaViola Jr, & Simpson, 2000). Analyses similar to those performed to compare Baseline with IMAGE-Desktop were conducted but this time for comparing IMAGE-Desktop and IMAGE-CAVE conditions. No differences on the measures of comprehension between those two groups were statistically significant. The report on the experiment (Gagnon et al., 2012) can be consulted for more details. Many reasons may explain the lack of benefits of immersion. First, IMAGE-CAVE did not offer a full user-centric interactive experience (Kerren, Ebert, & Meyer, 2007); IMAGE-CAVE provided a stereoscopic version of the 3D representation and the control of the Eye-Sys visualization tool via the Intersense Wand. Experimental results corroborate that this approach is not sufficient to get an improvement, at least for the kind of tasks participants had to achieve. Also problematic was the lack of responsiveness of the simulation tool (Multichronia) on the tablet PC due to a bug that appeared early in the experiment but was detected only at the end. It may have impaired the comprehension-related tasks (frustration, longer task completion time, etc.).

8.3 Focus time results on modules and tools

Baseline – Module versus tools

Figure 55 shows the focus time in percentage by module (I-REP, I-SIM and I-EXP) during a Baseline session while Figure 56 focuses on percentage by tool. I-EXP module is composed of

Microsoft Excel and a light version of Eye-Sys tool. Microsoft Excel is used for both modes (Tactical and SOP), while Eye-Sys is used only at the Tactical mode. Except for the SOP cycle 1 session, I-EXP module is the most exploited and its elapsed time generally increases with time. Participants seem to rely on I-EXP module in order to build their mental representation of presented scenarios. This hypothesis is supported by the fact that I-SIM and I-REP modules are respectively limited to generate runs (Multichronia), and to take notes (Microsoft Word). I-EXP module could count on Microsoft Excel to provide powerful analysis functionalities for complex understanding tasks.

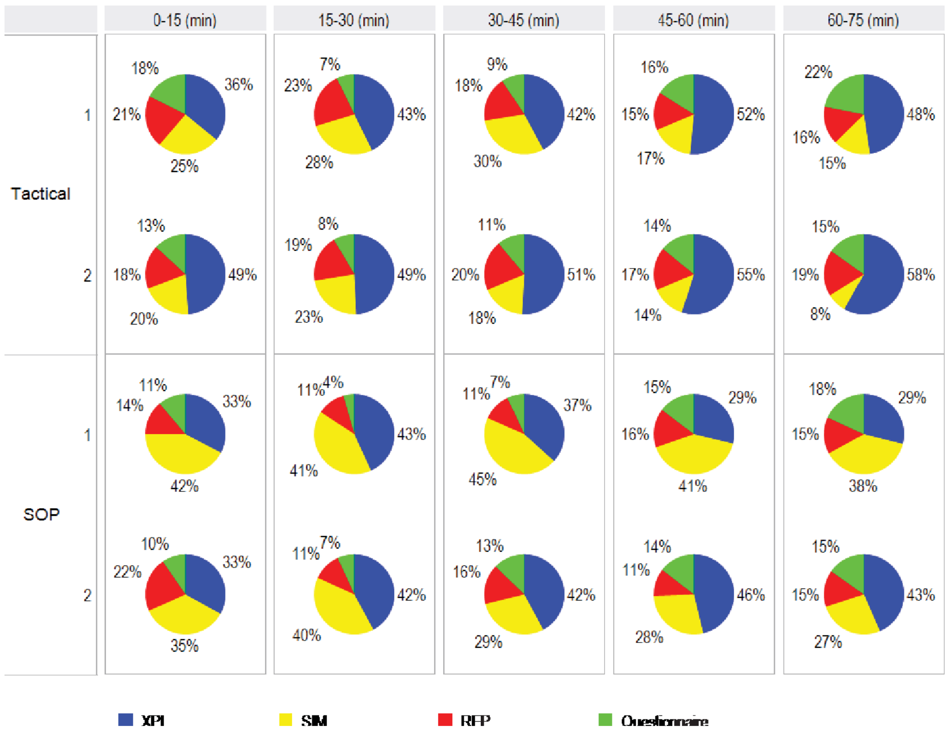


Figure 55: Baseline – Global usage by module.

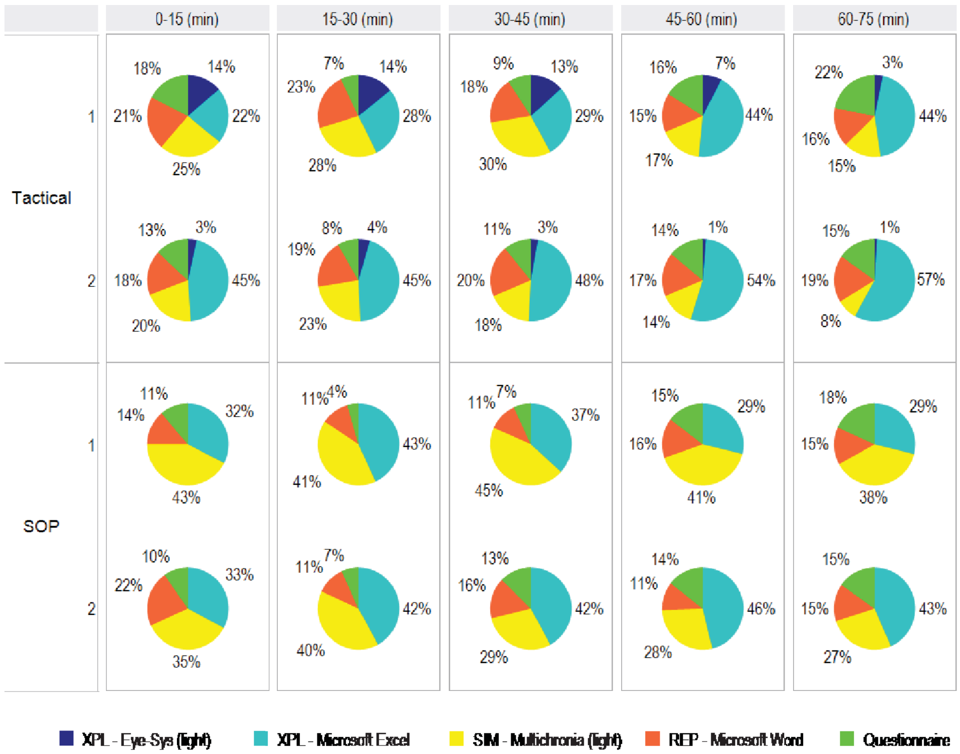


Figure 56: Baseline – Global usage by tool.

IMAGE-Desktop – Module versus tools

Figure 57 shows the focus time in percentage by module (I-REP, I-SIM and I-EXP) during a Desktop session while Figure 58 focuses on percentage by tool. Microsoft Excel is replaced by Eye-Sys and Tableau tools for I-EXP module, while CoGUI joins Microsoft Word for both modes of I-REP module. Finally, Multichronia benefits of simulation steering functionalities that supports the parameters mining process of the scenario. I-EXP module is mainly exploited for Tactical mode. Once again, participants seem to rely on I-EXP module in order to build their mental representation of the scenario. However, Multichronia plays a larger role for SOP mode, while simulation steering functionalities captures nearly half of session elapsed time. At last, a significant increase in time associated to I-REP module is observed at the end of each session. Participants take time to summarize their scenario understanding before sessions end.

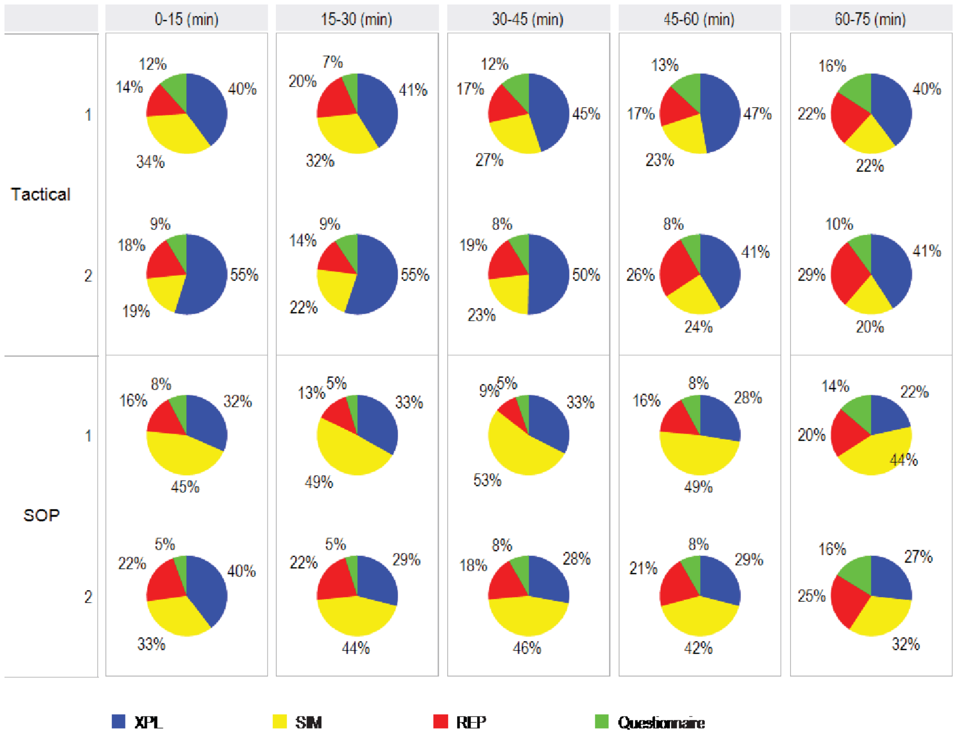


Figure 57: IMAGE-Desktop – Global usage by module.

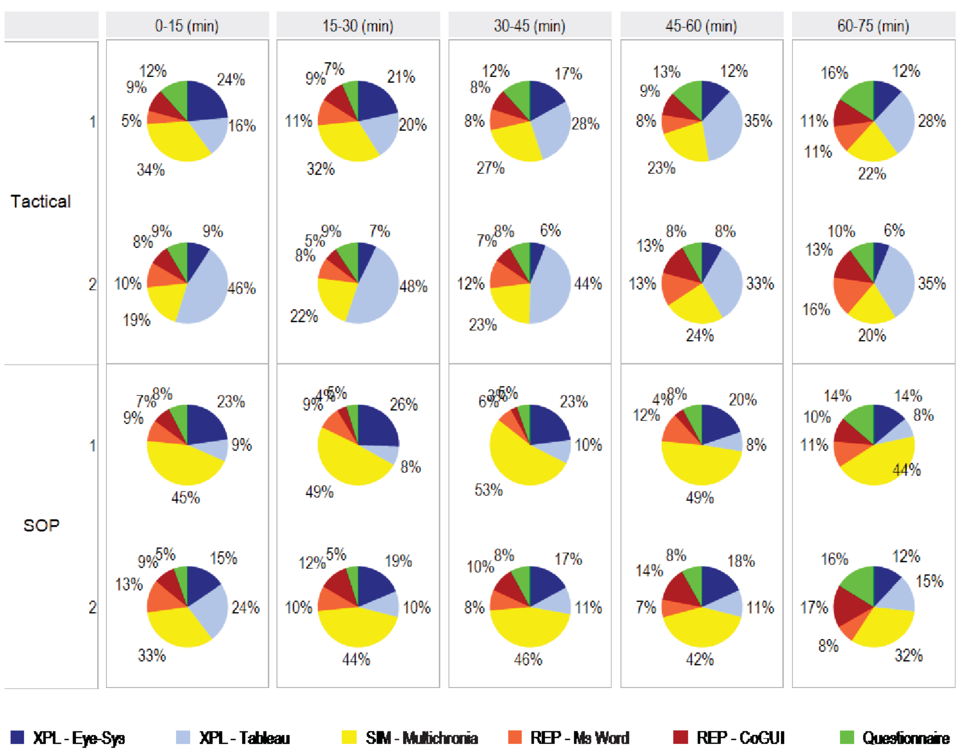


Figure 58: IMAGE-Desktop – Global usage by tool.

IMAGE-CAVE – Module versus tools

Figure 59 shows the distribution of participant keyboard and mouse events by modules and by tools over CAVE sessions. Because these sessions were not bio-monitored with any head or eye tracking systems, Figure 60 do not provide an accurate representation of focused time by modules and by tools. In addition, unmonitored changes were applied to the software tracking system, which makes difficult to analyse sessions among them. Nevertheless, intensive exploitation of I-EXP module for Tactical sessions and increase of use of I-REP module at the end of sessions confirm observed trends in Desktop sessions.

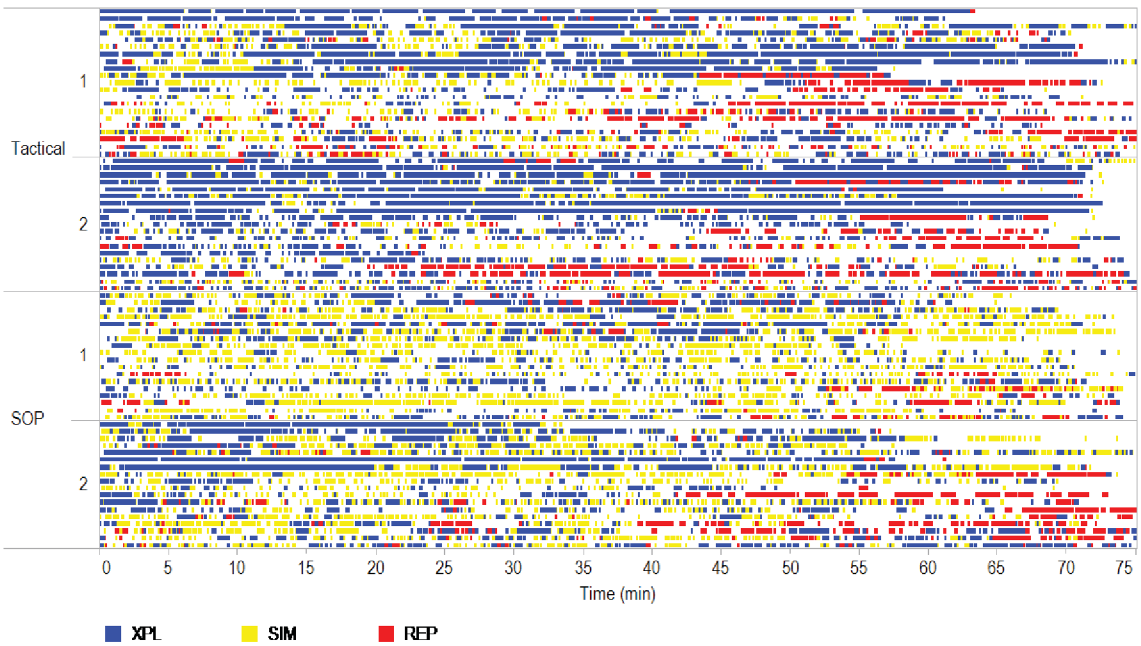


Figure 59: IMAGE-CAVE - Keyboard and mouse Events by module.

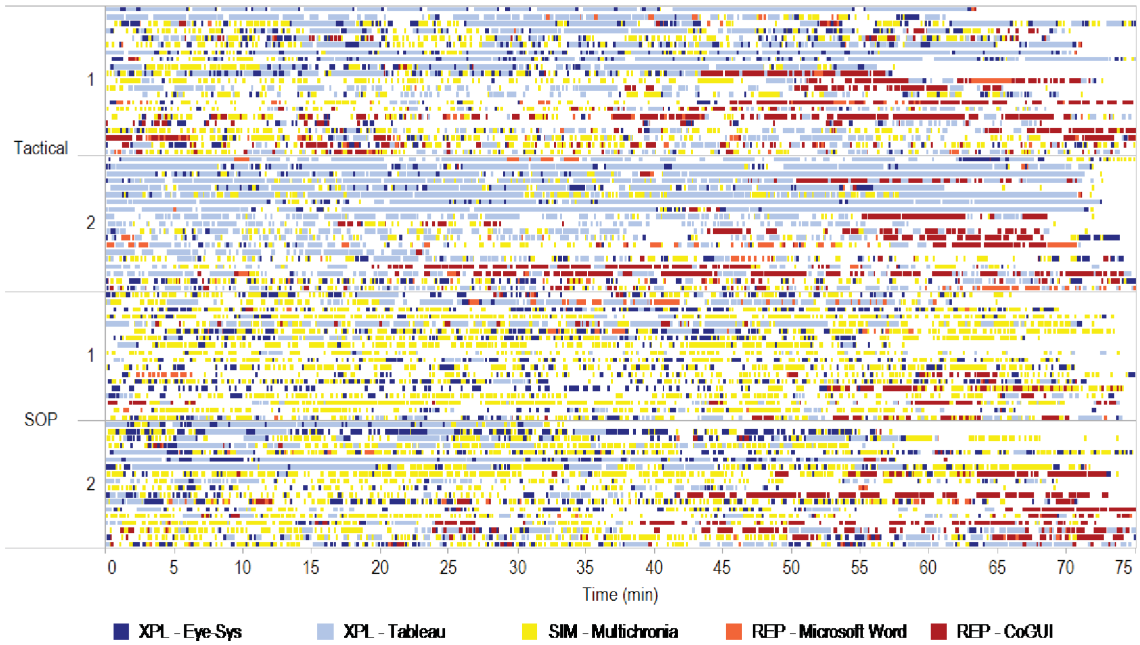


Figure 60: IMAGE-CAVE – Keyboard and mouse Events by tool.

8.4 Representation module (I-REP V1) results

As introduced previously, the I-REP concept has been experimented through two prototypes: I-REP V1, supporting individual understanding, and I-REP V2, adding support for collaborative understanding. Only I-REP V1 was used during the IMAGE V1 cognitive experiment enabling the user to:

- consult predefined graphs which represent the scenarios and examples of expected deliverables;
- insert, modify and delete vocabulary items (concept types and relation types);
- change the vocabulary hierarchies;
- insert, modify and delete conceptual graphs;
- change graph contents by inserting, modifying and deleting graph items: concepts, concept-relation links, concept-concept co-references, and nested graphs; and
- consult simulation graphs representing inputs and outputs of a I-SIM simulation that are dynamically generated by I-REP V1.

It did not include unification of a graph set and filtering the resulting view content and other features implemented in I-REP V2 for the collaborative understanding. In addition, a lot of available features were more or less required during experimentation: nested graphs and prototypics facilitating graph construction, vocabulary development and validation, and graph queries.

For the sake of the data analysis, I-REP V1 recorded a log entry for each significant application event e.g. the participant opens a graph, inserts a concept or the application generates a new simulation graph. The current section presents the most interesting findings from the various data collected in this huge log. The technical memorandum *IMAGE Representation Module - Experimental results* to be published in 2013 will report more detailed results on which these findings are founded.

8.4.1 Inadequate features for the experimentation

Vocabulary Edition is not useful once the problem is stated.

Participants were all trained to modify the vocabulary but did not use this capability. Potential explanations are that they did not see the need, the vocabulary was satisfactory and the time did not allow performing such edition. Examining user comments about the vocabulary, one can conclude that they did not see the plus-value: they mentioned that the vocabulary needed improvement while no one has tried to correct it. However, the kind of improvements mentioned was not regarding new vocabulary items but improvement to existing ones such as mismatch of terms between IMAGE tools and better documentation of terms. Although, this vocabulary edition was available, it is not such a surprise because the problem was already well stated and a good terminology was given to the participants. Actually, this vocabulary was elaborated by the research team while creating the scenario and defining the problem.

An important part of the IMAGE concept consists in defining the problem but it was excluded from the Image V1 scientific experimentation for a question of resources. For instance, the time, level of effort and expertise required for building a comprehensive experimentation with the right metrics was not available. Moreover the level of effort to run such experimentation was far away from the project capacity. Furthermore, the needs addressed by the vocabulary within the IMAGE concept are more specifically to address the collaborative work of people from different disciplines. Such a collaborative scientific experimentation is almost unfeasible. Means such as case studies or workshops are more realistic to evaluate the usefulness of the vocabulary piece.

The elaboration of a comprehension model has no plus-value once the problem is well defined.

As indicated in Chapter 4, a comprehension model is composed of a vocabulary (concept types and relation types) and a set of graphs. Therefore, the elaboration of a comprehension model means to create and update these elements. Vocabulary Edition was not used but what about the Graph Edition?

The results undoubtedly show that the creation of and modifications to graphs in this experimentation was not useful: (1) there was no correlation of the number of changes to graph with performance (Gagnon et al., 2012); (2) changes to graphs are decreasing with time; and (3) participants commented that Representation was not helping to answer analysis session questions and there was no need to formalise their comprehension or to look at the graphs since they had their mental model in mind during analysis. Finally, participants had likely enough time during dedicated sessions since work on the comprehension model decreased. It was stable (or slightly increasing) for Analysis sessions but it does likely not indicate usefulness but it is more simply related to the request to deliver the graphs.

Figure 61 shows that the evolution of I-REP V1 usage decreases from Cycle 1 to Cycle 2 and that it is mainly due to the important decrease of the Change Graph events and particularly for the Tactical scenario. Results showed a systematic increase in mean number of actions across cycles $F(1,166) = 14.421$, $p < .001$ and the SOP $F(1,50) = 33.698$, $p < .001$ levels. The analyses revealed no interaction between cycles and experimental conditions for both tactical $F(2,50) < 1$, N.S. and SOP $F(2,50) = 1.161$, N.S. levels.

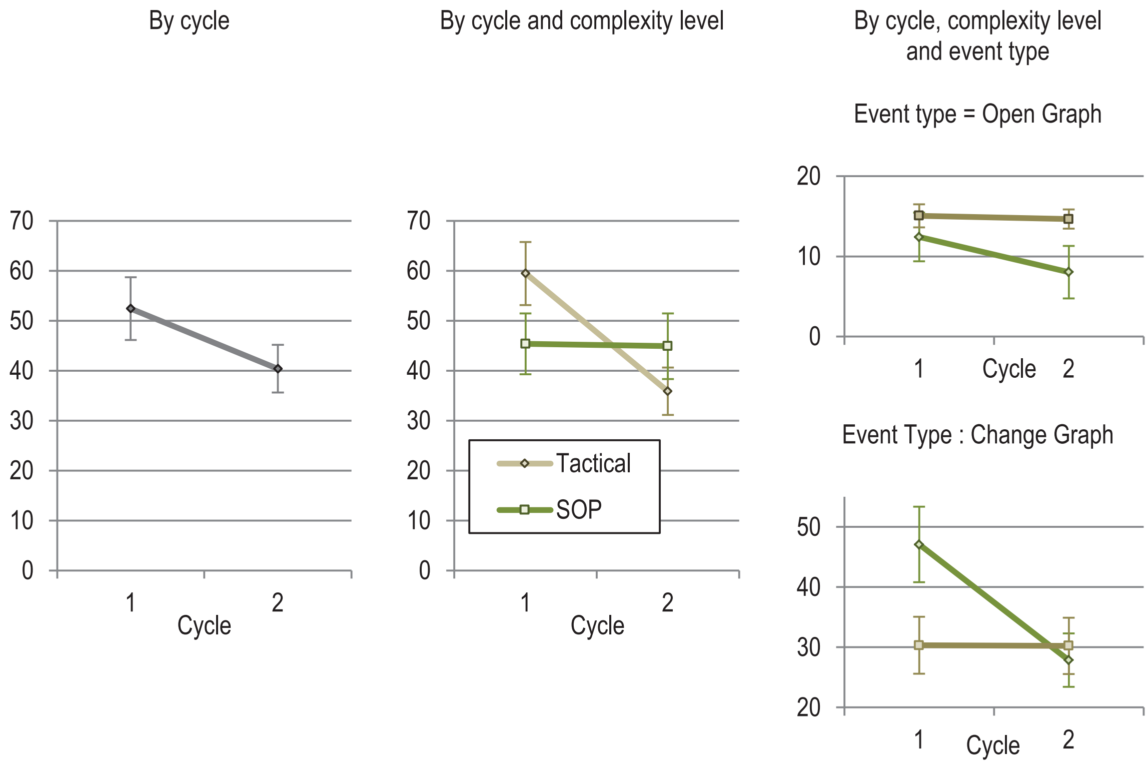


Figure 61: Average number of events by Cycle and Complexity Level (Tactical and SOP).

However, this experimentation was not well adapted to benefit from the creation of a comprehension model. In the context of this experimentation, a participant starts with a pretty good problem definition: the vocabulary is already well defined, variables are identified and some relations between the few variables are intuitively known e.g. more armoured on the side means less damage by RPGs. As explained, in a previous chapter introducing the scenario, it was required to proceed that way: the situation must be possible “to understand” within a few hours of IMAGE use and the situation needed to be intellectually interesting and challenging for the participant. Otherwise, either the experimentation time per participant would have been unacceptable or there would have been very high risk of not getting any meaningful results.

As mentioned, this set-up led to non usage of many features that would have been useful to sort out a CS from scratch. In addition, but at a lesser degree, the lack of friendliness and occasional lack of robustness reported by participants were tedious. Finally, it was impossible within the available time to train adequately all I-REP V1 features and enabling participant to master the creation of conceptual graphs. Such a tool can be powerful but requires a lot more hands-on time to exploit this power. Simple drawing was not an issue but knowing what are useful usages of conceptual graphs do not happen overnight and can hardly be overcome in an experimentation setup. Individual methods to fully exploit specialized tools are developed through various experiences with the tools.

8.4.2 Useful features for the experimentation

Automatic or assisted generation of graphs is a useful approach.

User comments stating the usefulness of simulation graphs automatically generated by I-REP V1 are coherent with the results showing that they were more used than any other graph categories and their usage increase from Cycle 1 to Cycle 2 for the SOP scenario level. Believing that assisted creation of graphs would also be appreciated is a common sense extrapolation. Some features were available but participants were not trained to do it and the experimental scenario did not really require to explicit a mental model.

Figure 62 shows that Simulation graphs were the most frequently used followed by Deliverable, Scenario and Others.

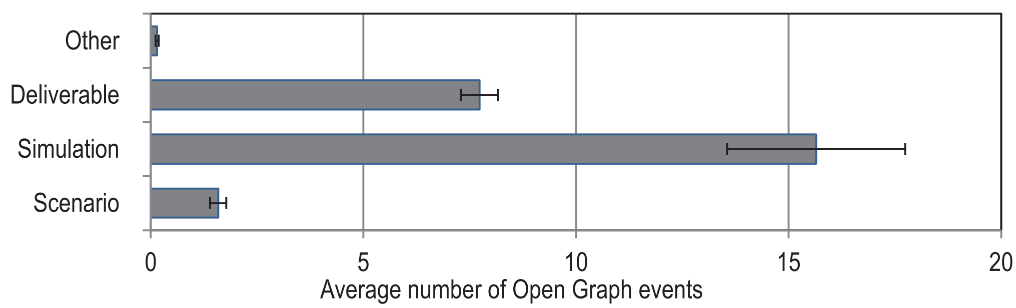


Figure 62: Average number of Open Graph events by Graph Types.

More complex a situation is, more Representation artefacts are useful

The SOP level scenario was more challenging than the Tactical level. Results demonstrate that: (1) SOP Simulation graphs consultation has increased from Cycle 1 to Cycle 2 (while Tactical Simulation graphs usage has decreased); and (2) that SOP Simulation graphs are more consulted during SOP Representation sessions than during SOP Analysis sessions (while it is the opposite for Tactical Simulation graphs). The increase or decrease between from Cycle 1 to Cycle 2 is a good indicator of usefulness. In addition, participants had more time in dedicated Representation session than during Analysis session to examine graphs. Consulting more frequently during Representation is also a good indicator of usefulness.

A wider screen tends to improve work intensity on graphs.

The CAVE wide screen used for I-REP does not show more information than the Desktop screen since they have similar resolutions. As shown in Figure 63, results indicate that the average number of Change Graph events is higher for IMAGE-Desktop than for IMAGE-CAVE. However, these are not statistically significant and do not allow to conclude that a wider screen leads to more intensive work on a comprehension model for both Open Graph ($t(41) = .83$, $p = .41$) and Change Graph ($t(41) = 1.95$, $p = .058$).

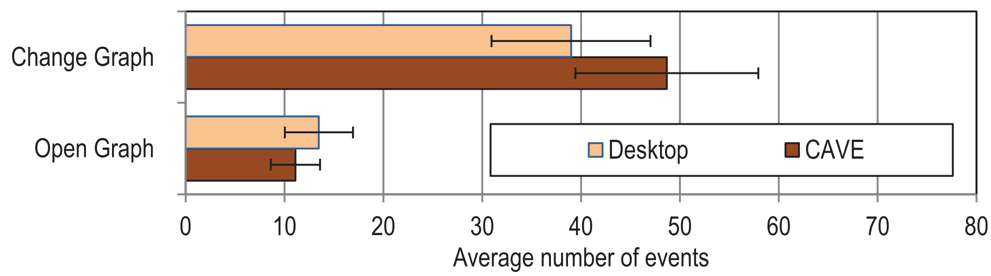


Figure 63: Average number of Events by Event Type and IMAGE Condition (Desktop and CAVE).

Perception of participant does not tend to fit their usage of Representation for a higher complexity level.

Figure 64 shows an increase of average Open events for the SOP Scenario Level (higher complexity). It also show that SOP Simulation graphs are even more used during Representation Sessions than during Analysis Sessions whereas it is the opposite for Tactical simulations i.e. Tactical Simulation graphs are less used during Representation Sessions.

For the SOP Scenario Level, these results are in contradiction with a strongly shared comment among many participants stating I-REP was not useful for analysis. Actually, on average participant opened graph as often for Representation sessions as they did for Analysis session. However, the Tactical and SOP Simulation graph Categories were the most frequently used and, in the mind of participants, they are visibly less associated to representation than they are to simulation. Although these graphs do not result from a participant representation effort, they are actually generated by I-REP and are representation artefacts.

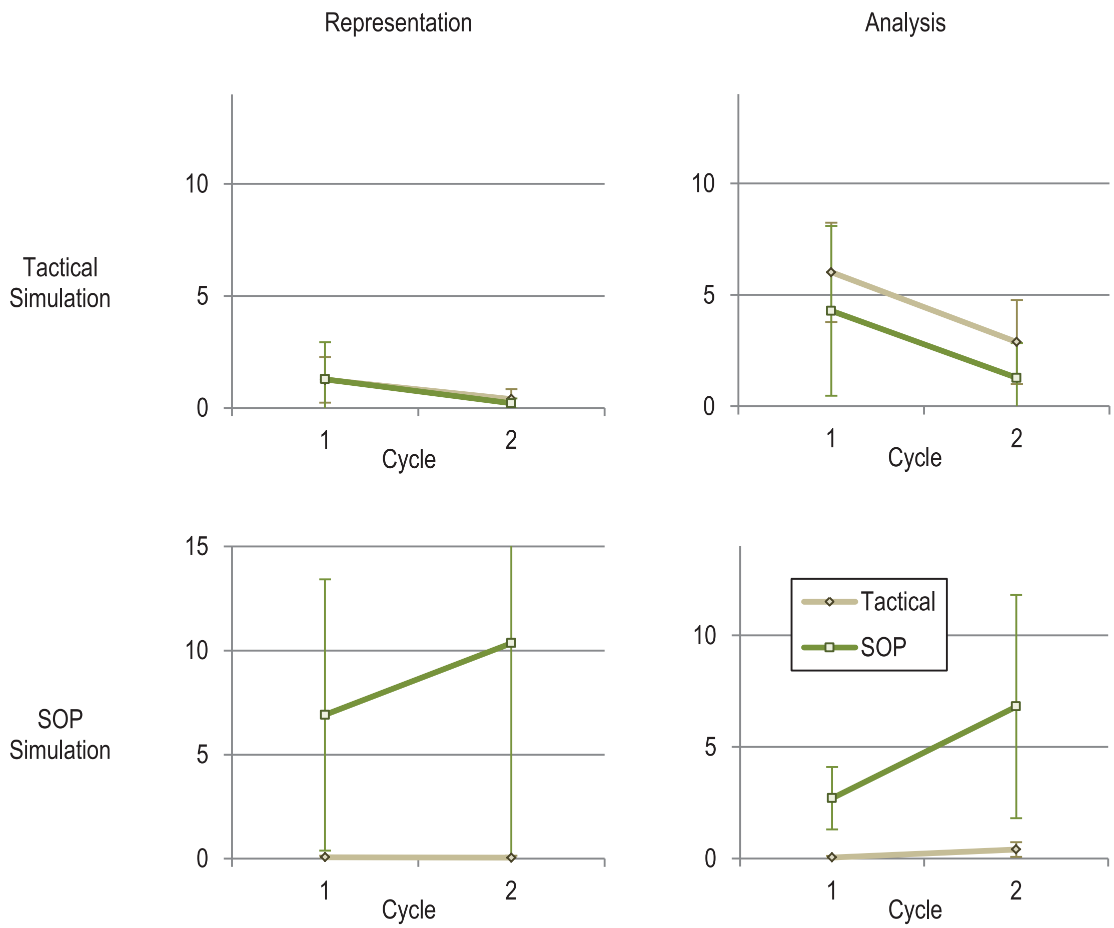


Figure 64: Average number of "Open Graph" by Graph Category (Representation and Analysis), Session Category, Scenario Level (Tactical and SOP) and Cycle

8.5 Simulation module (I-SIM) results

Multichronia provides a series of simulation-related functionalities that aim at supporting users in their activities used for exploring and comprehending CSs. The relevance and the efficacy of these activities were evaluated from the analysis of how participants used Multichronia during IMAGE experiment.

8.5.1 Multichronia usage

Figure 65 presents the average number of activities performed by participants during their two 75-minute sessions. For most activities, little difference was observed between Baseline, IMAGE-Desktop, and IMAGE-CAVE conditions. The exceptions were *Drag Node*, *Start Simulation*, *Change Param*, *Change MoP on Axis*, and *Create Tab*. More differences were observed between IMAGE-Desktop and IMAGE-CAVE than between IMAGE-Desktop and Baseline, probably because Multichronia was implemented on a Tablet PC in IMAGE-CAVE. Future usability and user-friendliness improvements should focus on most frequent activities, namely *Select Branch*, *Drag Node* and *Create Branch*.

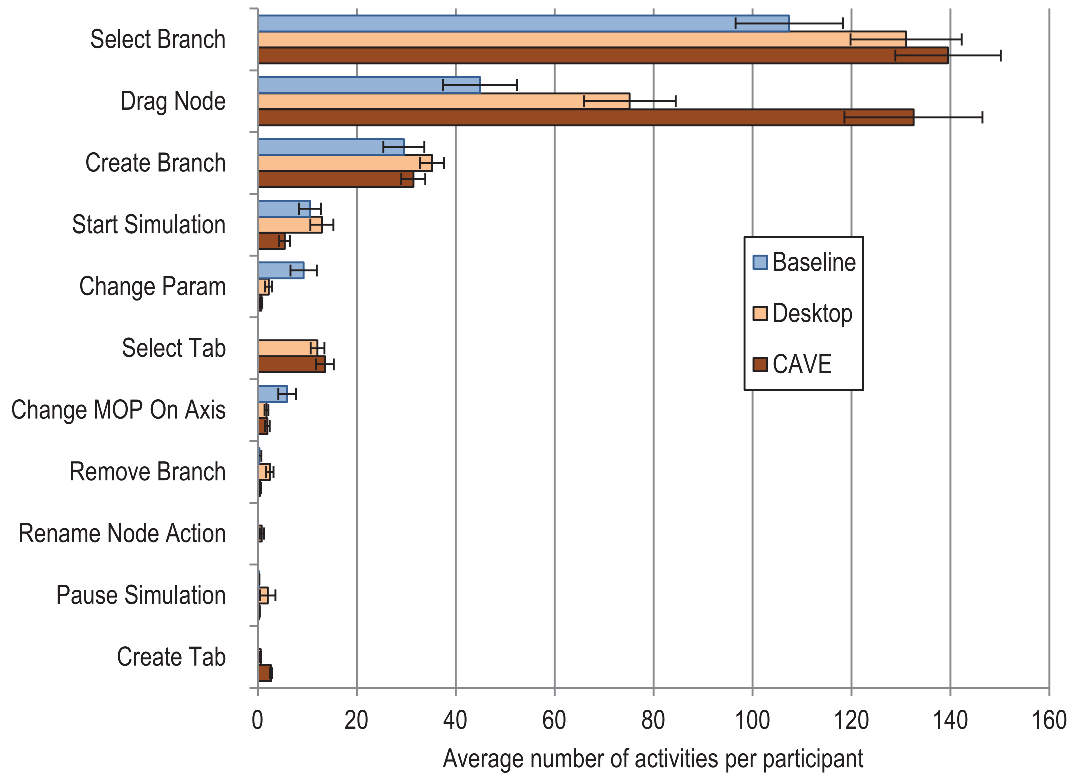


Figure 65: Average number of activities performed by each participant with Multichronia for each condition (error bars represent standard error).

8.5.2 Temporal profile

Figure 66 shows the number of simulation runs per participant created for each 10-minutes time window. This granularity was chosen to even out small variations. On average, the number of runs created by participants was small at the beginning, partially due to the first one or two minutes participants took to read questions and to become fully operational. Then, participants created more simulations in the following 10-minute block and the next one, where it reached its maximal value. It finally decreased slowly until the end. The most interesting finding was that users of the three groups (dashed lines) created simulation runs according to this pattern for both cycles (blue and red).

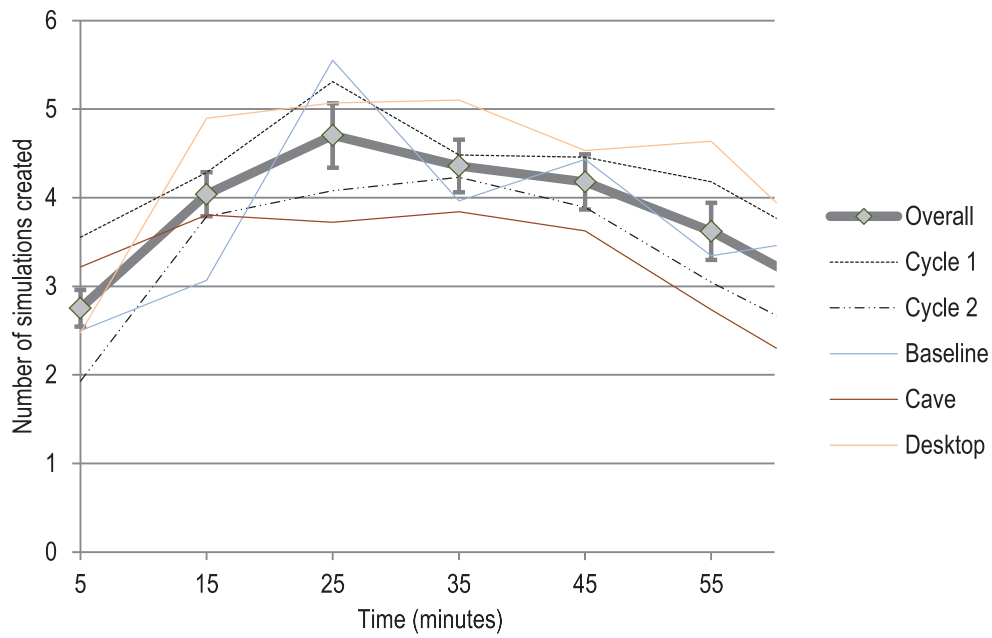


Figure 66: Average number of simulations created by participants for each 10 minute-period (from -5 to +5 minutes; error bars represent .95-confidence intervals).

8.5.3 Parameter exploration strategies

Participants were free to explore the parameter space in order to get a better understanding of the effect such values on the simulation outcomes. Figure 67 shows the coverage of the parameter space at the Tactical level with values between 0 and 10 for UAT, SAT, and 0 to 5 for NIED, for a total of 726 combinations. The area of each circle is proportional to the total number of simulations instantiated by all participants according to these combinations of parameters. Participants focused their exploration of the parameter space on specific areas, namely the most extremes values, the “periphery”, and the diagonal corresponding to pairs of identical values (UAT, SAT).

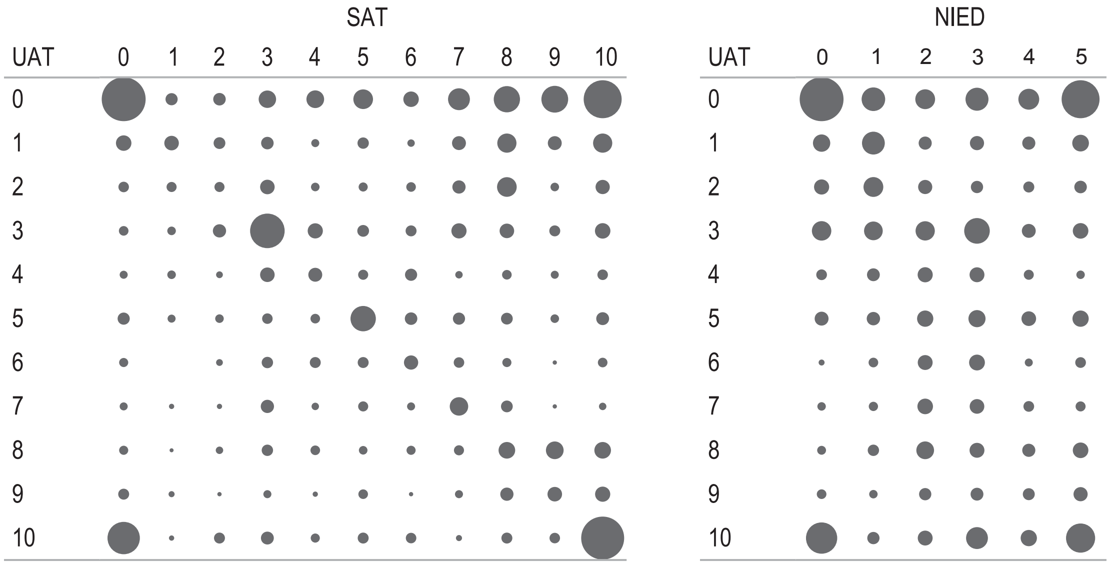


Figure 67: Distribution of the combinations of parameters tested by all participants.

Participants of the IMAGE condition were provided, among other things, a visual tool that displayed continuously the parameter space explored (see Figure 36 item D). Restricting the use of specific simulation and visualization tools to each group offers the opportunity to use similarities and differences in participant exploration behaviours to discriminate the influence of the tools from the inherent strategies humans employ. Both similarities and differences were observed from the experimentation.

A noteworthy similarity is represented in the left-hand side of Figure 68 that shows a high proportion (about 50%) of identical parameters tested by participants from Baseline and IMAGE conditions for most frequent encountered combinations of parameters. For instance, the most common combinations or parameters (0, 0, 5) was the same in both conditions, hence a value of commonality of 100% at .1% (1 / 726) of the parameter space. If a higher (1.8%) but still small proportion of the parameter space is taken into consideration, 7 out of 13 (54%) most frequent parameters explored were the same for both groups. By contrast, the random selection of parameters by participants would have resulted in 3 out of 13 (21%) identical combinations of parameters. Overall, 64% of the parameter space between both groups was identical. Were much more time to be available to explore or more participants to be included in the experimentation, the level of commonality would have reached 100% because the parameter space would have been explored utterly. Also similar was the temporal evolution of the parameter space coverage (i.e. fill-factor vs. time). The right-hand side of Figure 68 shows the average of the percentages of all participants' parameter space coverage according to exploration time. Two outliers, one of each group, were excluded because they created four to five times the average number of simulations. On average, participants from IMAGE and Baseline conditions explored the parameters at the same and constant pace (error bars represent .95 confidence intervals). The confidence interval is smaller for IMAGE because of a higher (2.8×) number of participants for this group. Participants in IMAGE and Baseline conditions recreated respectively 14.6% and 14.7% simulations more than one time hence a loss of productivity. No statistically difference was observed between IMAGE and Baseline conditions $t(56) = .012$, p = N.S.

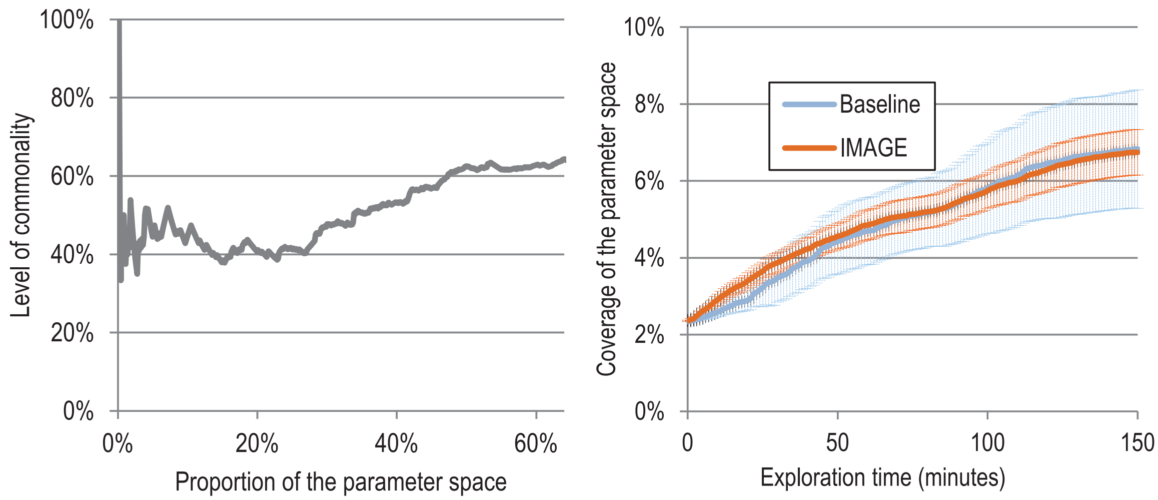


Figure 68: Proportion of common parameters (UAT, SAT, NIED) simulated both by Baseline and IMAGE conditions (left) and average coverage of the parameters space over time (right).

The discrepancy (Fang, Li, & Sudjianto, 2005) of the combinations of variables simulated by participants is a measure of the ability to efficiently get the global picture of the model, i.e. avoid over testing the model under the same conditions while neglecting other conditions. A complementary measure is the coverage that assesses how much the sampling covers the parameter space entirely, in equally manner and isotropically distributed (Gunzburer & Burkdart, 2004). Both measures were calculated at the Tactical level for both cycles and three conditions as illustrated in Table 7. A one-way repeated measure ANOVA revealed a significant difference between IMAGE and Baseline when comparing their discrepancies, $F(1, 112) = 5.7$, $p < .019$, but not when comparing their coverage, $F(1,112) = .971$, N.S.. The effect of the cycle was statistically significant for the discrepancy measure, $F(1, 112) = 348$, $p < .001$, and the coverage measure, $F(1, 112) = 30.9$, $p < .001$. Considering the small difference in discrepancy between IMAGE and Baseline, we can conclude IMAGE supports or enforces only by a small margin a more uniform space exploration behaviour than Baseline does.

Table 7: Average discrepancy and coverage of participants at the Tactical level by cycle and condition (standard error).

	Cycle 1		Cycle 2		AVG (cycle)	
Condition	Discrepancy	Coverage	Discrepancy	Coverage	Discrepancy	Coverage
IMAGE	.015 (.000)	.915 (.042)	.012 (.000)	.818 (.011)	.013 (.000)	.866 (.012)
Baseline	.015 (.000)	.931 (.018)	.011 (.000)	.777 (.023)	.013 (.000)	.854 (.028)

The only difference observed between IMAGE and Baseline was that the 10% most frequent combinations of parameters represents 49% of all simulations instantiated in the Baseline condition while it represents 57% in the IMAGE condition. This result suggests that participants in the IMAGE condition explored the parameter space in a more similar manner, probably as the result of the better feedback provided by a visualization tools.

8.5.4 Multichronia use and performance

Multichronia was the tool on which participants sent most of their time (31% on average). From a module perspective, I-SIM (Multichronia) was second with 31% of user attention versus 33% for I-EXP (Tableau and Eye-Sys). Regression analysis between pattern of use of Multichronia and comprehension performance were carried out, so future experiments could be designed to test if some Multichronia functionalities and processes can cause an improvement of CS understanding. Using more Multichronia was correlated (R^2 = .158, p = .003) with an improvement in understanding but the specific ways of using it still eludes us.

8.6 Exploration module (I-EXP V1) results

As introduced previously, the I-EXP concept has been experimented through two prototypes: I-EXP V1, supporting individual understanding, and I-EXP V2, adding support for immersive virtual exploration and collaboration aspects. Only I-EXP V1 (I-EXP V1 Tactical and I-EXP V1 SOP) was used during the IMAGE V1 cognitive experiment. The technical memorandum *IMAGE Exploration Module - Experimental results* to be published in 2013 will report more detailed results.

8.6.1 IMAGE-Desktop – Temporal profile of the tools

Figure 69 compares the elapsed time in percentage among I-EXP tools during Tactical Desktop sessions. The Tactical version of Eye-Sys tool provides a geospatial view that supports the understanding of the scenario. Participants exploit intensively the geospatial view in the first minutes of cycle 1 sessions, and then gradually shift their attention on Tableau in order to further analyse generated data.

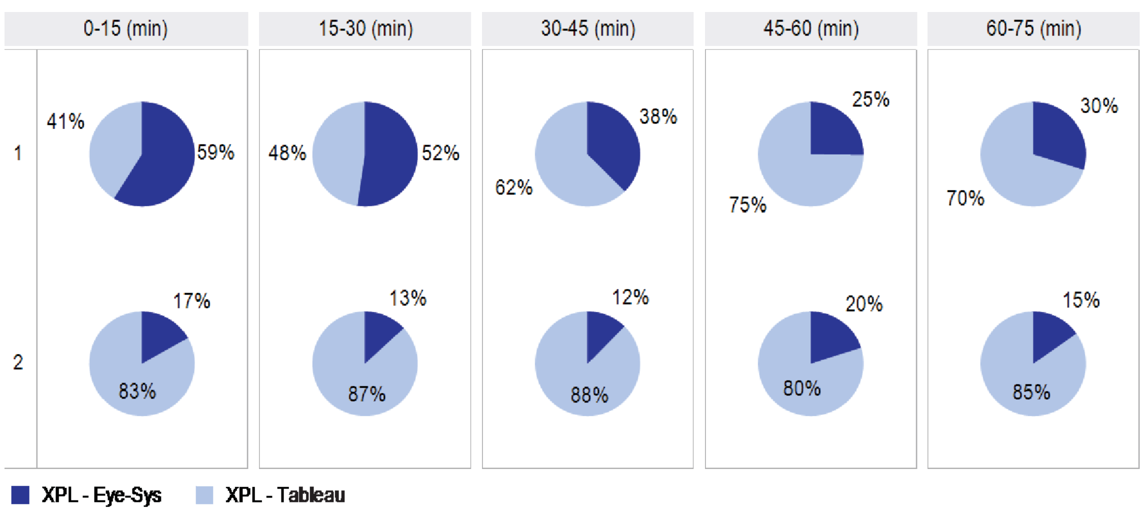


Figure 69: Usage of the I-EXP desktop tools at the Tactical level.

Figure 70 compares the elapsed time in percentage among I-EXP tools during SOP Desktop sessions. Eye-Sys provides dedicated views designed to support the scenario analysis. Similar features to those offered by Tableau help Eye-Sys to keep the major part of participant interest for the use of I-EXP SOP tools.

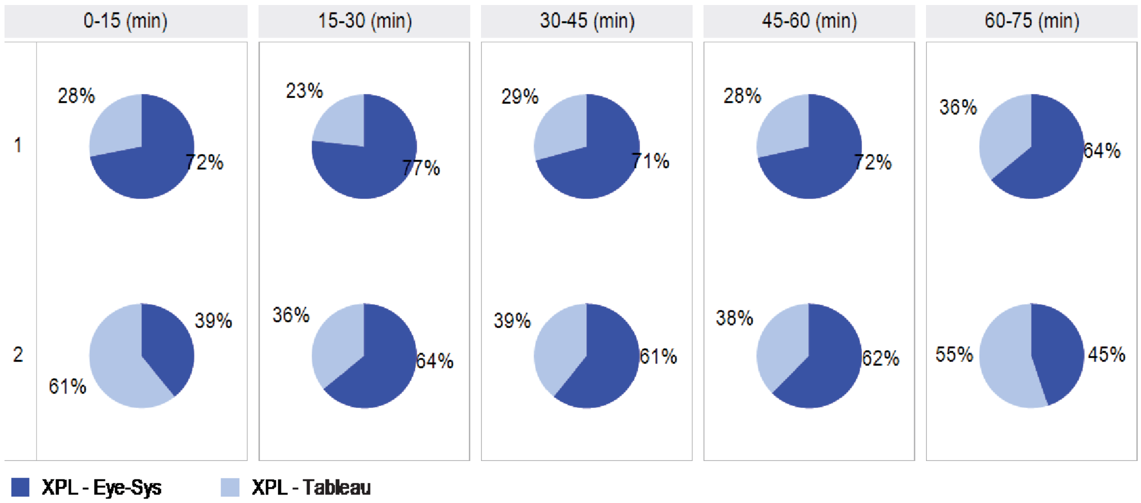


Figure 70: Usage of the I-EXP desktop tools at the SOP level.

8.6.2 IMAGE-CAVE – Temporal profile

Figure 71 shows the distribution of participant actions for I-EXP tools over CAVE sessions. Because these sessions were not bio-monitored with any head or eye tracking systems, the figure does not provide an accurate representation of focused time by tools. Nevertheless, intensive exploitation of Tableau for Tactical sessions and increasing of use of Eye-Sys for SOP sessions confirm observed trends for Desktop sessions.

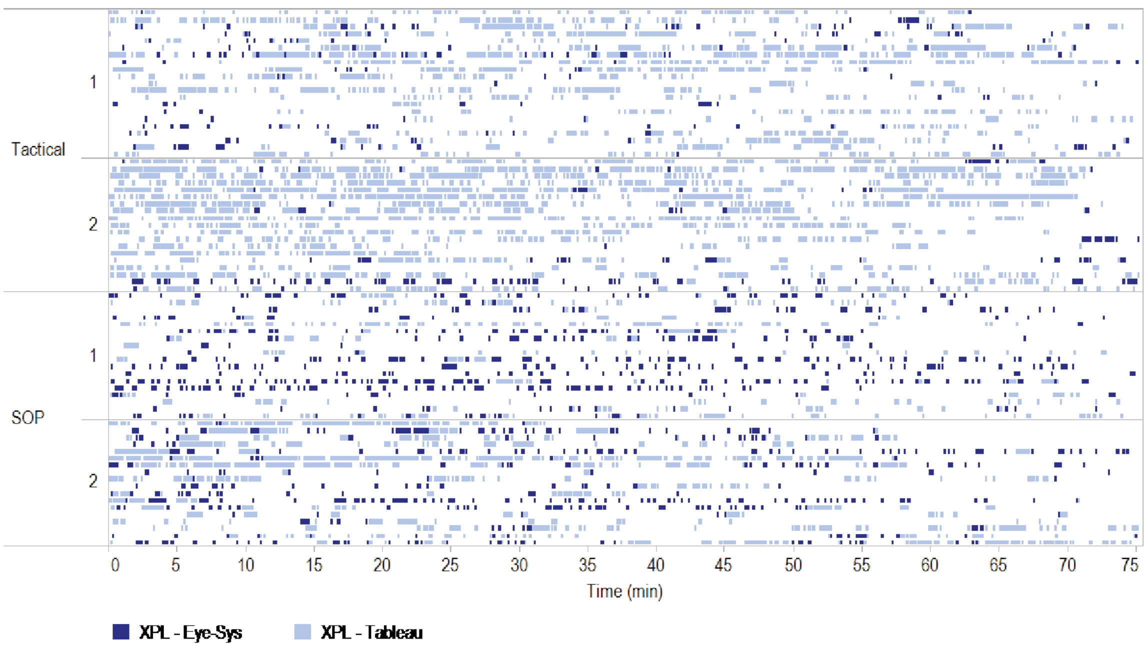


Figure 71: IMAGE-CAVE - Temporal profile of the usage of the I-EXP V1 tools (at Tactical and SOP levels).

Figure 72 shows the distribution of participant actions for I-EXP tools over CAVE sessions. CAVE actions conducted in 3D immersive mode versus non-immersive mode are distinguished. The graphics reveals that the 3D CAVE mode was underexploited.

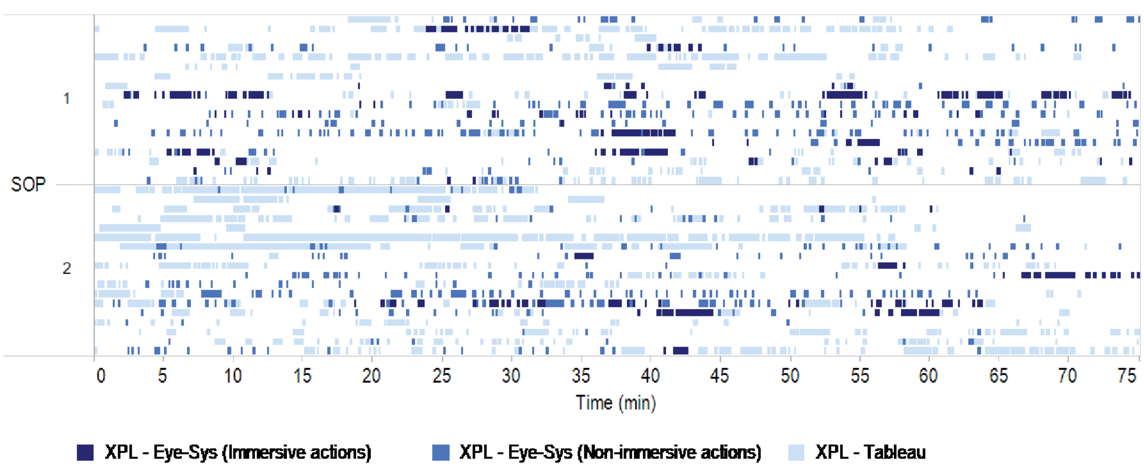


Figure 72: IMAGE-CAVE - Temporal profile of non-immersive and immersive actions logged during the experiments (at the SOP level).

8.6.3 Synergy aspect between the I-SIM and I-EXP V1 tools

Figure 73 shows two use patterns between I-EXP tools and the I-SIM tool. Eye-Sys and Multichronia present a higher level of synergy that optimizes the use of both tools. Eye-Sys reads

Multichronia data stream in real-time and automatically updates its views. Even if Tableau is virtually linked with Eye-Sys through a set of shell commands and pre-filled views, Tableau stays a stand-alone software requiring manual actions to get the latest view. These two styles could be observed into Figure 73 where Eye-Sys and Multichronia events are highly intertwined together, while Tableau and Multichronia events are grouped in separate blocks.

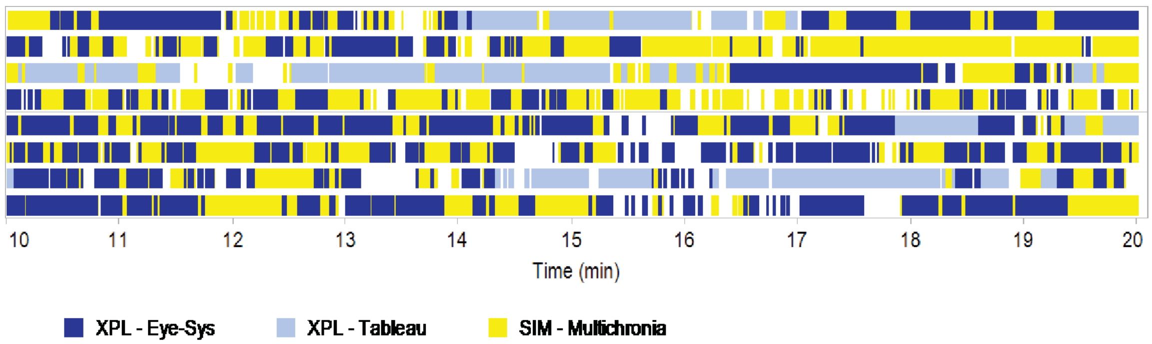


Figure 73: Examples of the synergy between the tools of the I-SIM and I-EXP V1 modules.

8.6.4 Desktop SOP tool as a branch starter in Multichronia

Figure 74 demonstrates how Eye-Sys tool could help the participant to optimize his data mining strategy. Because Eye-Sys and Multichronia are temporarily synchronized, the “Performance totale par branche” graphic could be used for positioning its time indicator on the optimal performance value to branch a simulation run.

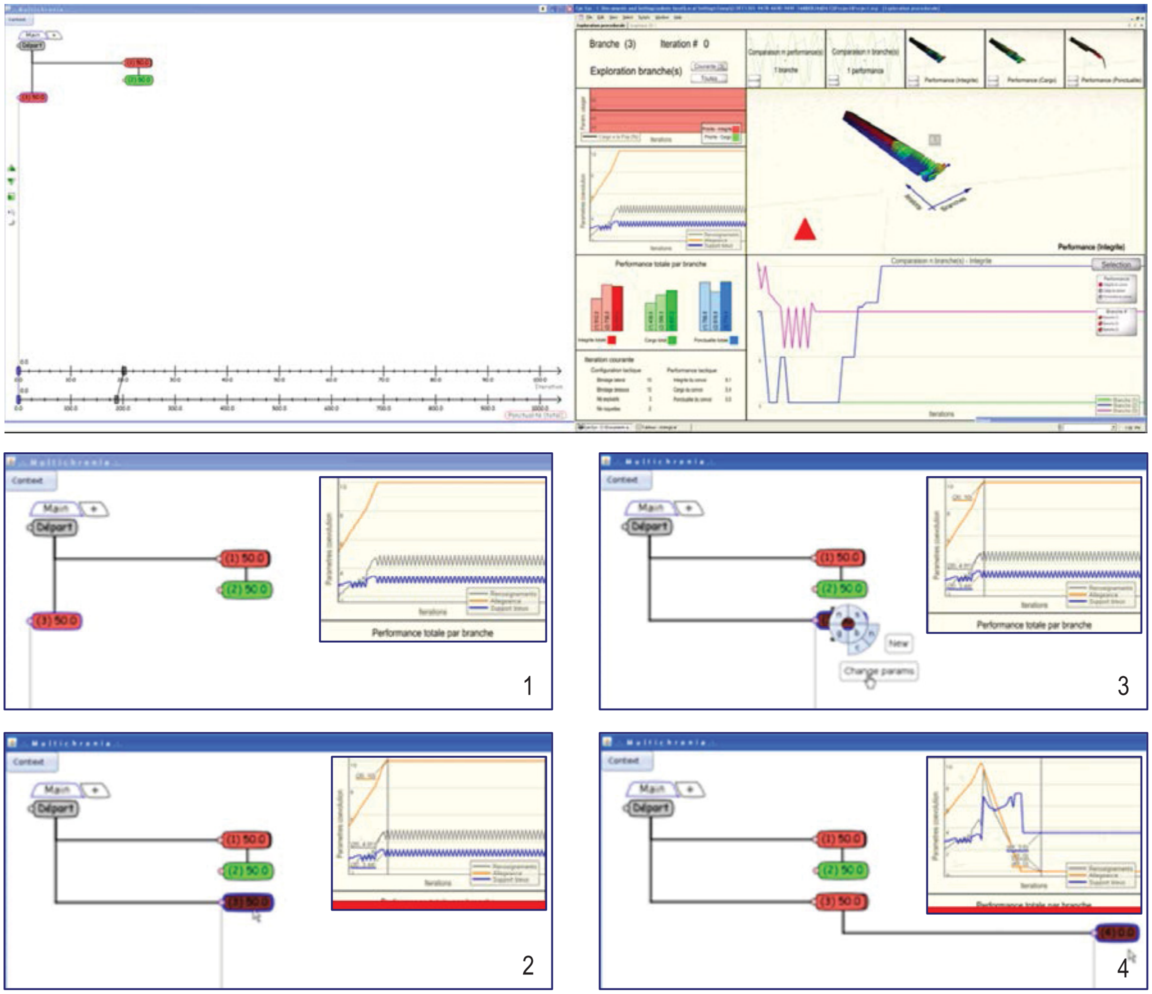


Figure 74: The I-EXP V1 SOP tool used as a starter of branch in Multichronia, the I-SIM tool.

8.6.5 Parameter exploitation strategies

Figure 75 illustrates how many times each simulation was generated by each session type (Baseline, IMAGE-Desktop and IMAGE-CAVE) and mode (Tactical and SOP). The figure clearly shows density differences between session types and modes. Logically, SOP mode generated more simulations than Tactical mode. And Desktop sessions generated slightly more simulations than CAVE version due to a friendly physical interface setup. IMAGE-CAVE setup suffers of a lack of usability generated among other by a tablet processing issue, a too wide field of view and a limited physical space for mouse movements. Finally, other patterns could be observed for Tactical mode where participant decided which simulation to generate and for SOP mode where algorithm has tendency to converge through a lower level of SAT and UAT values.

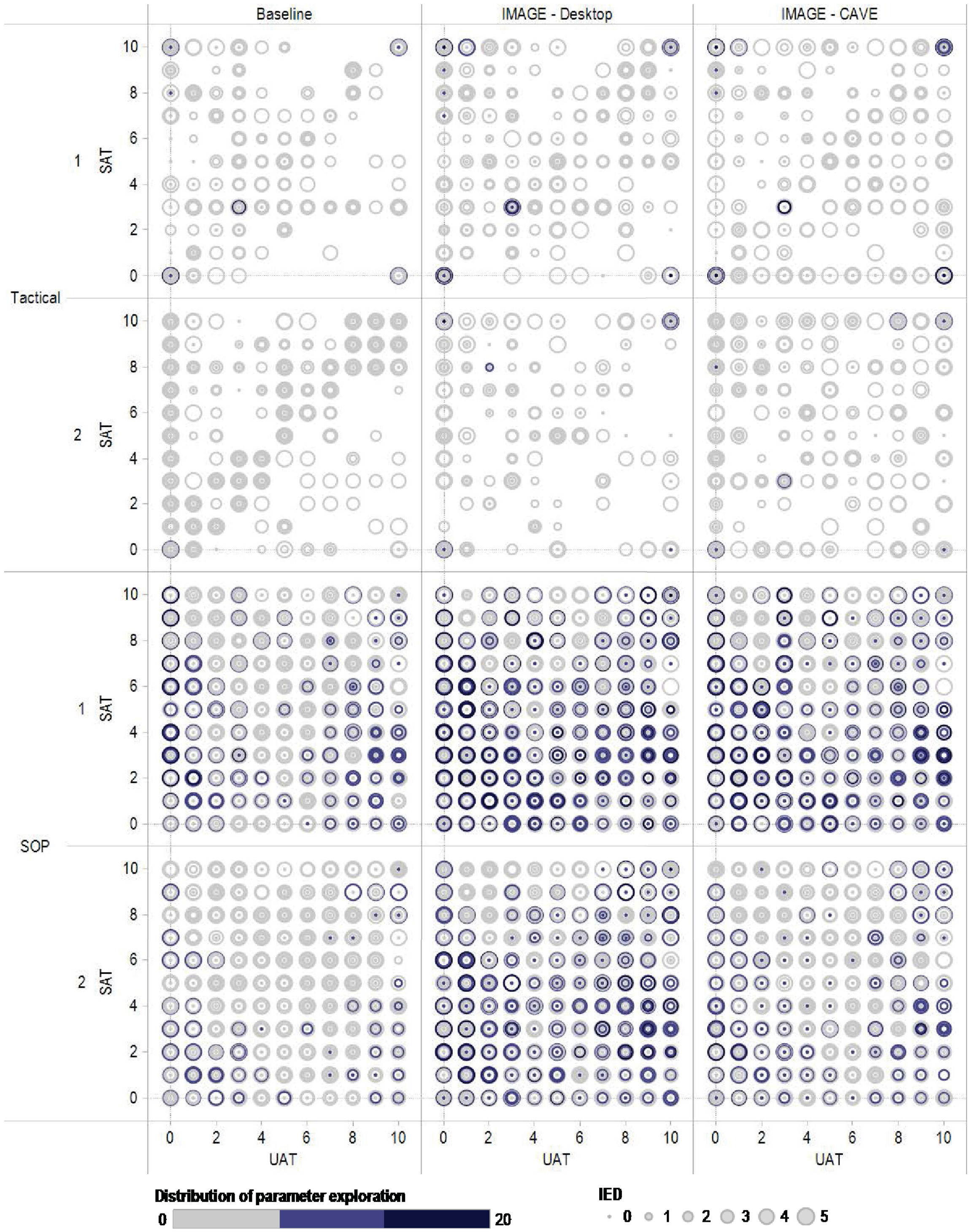


Figure 75: Spatial parameters distribution of the Tactical simulations in terms of SAT, UAT and NIED.

8.7 Summary

Table 8 summarizes the experimental results presented in Chapter 8. The benefits of IMAGE are obvious at a lower (but still challenging) level of complexity. Clearly, IMAGE failed to provide benefits at the higher (and maybe excessive) level of complexity. Overall, two out of three measures showed an improvement of comprehension.

Table 8: Summary of augmentation of comprehension provided by IMAGE.

	Low complexity (Tactical)	High complexity (SOP)	Overall
Comprehension: analysis	No*	No	Yes
Comprehension: prediction	Yes	No	No
Comprehension: mental models	Yes	No	N/A
Response selection fallacy	Yes	No	Yes
Immersion added-value	No	No	No

*Close to statistical significance

One possible cause of the failure of observing the benefit of IMAGE at the SOP level is the unrealistic ambitious target fixed in term of complexity (see Section 3.3) that was probably out of reach for a human considering the time and tools available. Were the tools adequate for the task? Was sufficient time given to comprehend the CS with the tools? For the first question, it seems still possible that better tools (i.e. IMAGE) could improve comprehension even at high level of complexity; on the average, participants in IMAGE group performed always better than Baseline in all measures, but none these differences were statistically significant. It suggests that, for future work, improving the tools is a better approach to get statistical significance than augmenting the number of participants (it would be impractical considering the time required to conduct such experiment). As for the second question, it is important to highlight the fact that the gap between Baseline and IMAGE at the SOP level, between Cycle 1 and Cycle 2, remained the same for analysis and RSF score while it decreased for prediction score. It is doubtful than more time dedicated to the analysis phase would have reversed current tendency and then provided additional benefit for IMAGE over Baseline. Experimental results do not support the added-value of immersion. Retrospectively, this finding is not surprising and future improvements, implemented in IMAGE V2, could have overcome limitations that caused the absence of gain of comprehension. Yet the benefits of immersion remain to be demonstrated.

Both similarities and differences were observed when comparing how the three groups used the tools. For instance, some functions were more popular in a given working environment than in another. On other aspects, like the parameter space exploration strategy and average coverage of the parameters space over time, all groups followed similar strategies. One difference between Baseline and IMAGE was on discrepancies of the sampling of the parameters space explored by participants.

8.8 Lessons-learned

Without a doubt, the IMAGE experimentation was extremely useful in terms of results but also in terms of learning for all team members. In addition to greatly increase their knowledge on complexity, cognitive psychology and engineering researchers have mutually learned issues of the other discipline. As mentioned previously, fifty-height people have undergone between nine (9) and fourteen hours (depending the group they were in) of effort distributed over several sessions. Many lessons can now be share about such an experimental context. Almost all of them are directly related to the objective of measuring the augmentation of comprehension combined to the exploratory nature of the project. This has required the development and adoption of methods that were hard to implement (Gagnon et al., 2012).

The following elements should be considered in conducting similar experimentation in the future:

- Commercial tools: Developing new software or adapting existing one while preparing experimentation is debatable. Development is always longer than expected and usability is rarely achieved in the first version of any software. IMAGE modules were not an exception. Maybe the concept IMAGE should have been first tested with commercial tools.
- Length: The amount and time spread over many experimental sessions required a lot of synchronization efforts. Being able to conduct the experiment within a day and increasing the number of participants that perform the experiment at the same time (more participants per experimenter) should be considered.
- Data collection: Large amounts of data were collected by various tools. Complicated processing was required to synchronize time used by the different data collection tools. Data analysis was difficult and the vast amount of data could not easily be processed by typical statistical software. Data had to be pre-processed in different ways throughout the evolution of the analysis. Reducing as much as possible the number of data collection tools and a priori synchronization of their clock would greatly reduce efforts.

Even considering all of these improvements, a second study of this sort would most likely not be as beneficial. Knowing the required time and effort to conduct such experimentation, less ambitious objectives should be evaluated. For instance, a case-study approach instead of a group comparison approach may be more appropriate.

This page intentionally left blank.

9 Conclusion

In this report, the IMAGE defence scientists have proposed a toolset concept aiming at disentangling CSs more quickly, specifically targeting the collaboration of experts from different disciplines trying to reach a shared understanding of a CS. The goal of the concept is two-fold: increasing understanding of a CS and enabling individuals to share their comprehension. The elaboration of the IMAGE concept was mainly guided by three principles or assumptions: (1) Iterative understanding: a common understanding is reached through revision and sharing of successive representations of the situation; (2) Human-driven toolset: the common understanding is above all a human task involving team members using tools their own way; and (3) Synergy of technologies: the common understanding is better served by the synergy of technologies than a single one. IMAGE researchers have identified capabilities that each of the various tools for representation, scenarization, simulation and exploration can bring.

The Representation module (I-REP) consists in expliciting a mental model into a comprehension model. But properly grasping what another individual means about a CS implies some challenges: enabling an individual to explicit thoughts; providing a team with means to develop and maintain a common terminology; and assisting in merging individual perspectives. Ideally, at the end of the collaborative effort, the resulting common model of a CS should not include any redundancy or contradiction. On the other hand, pieces created by various team members should complement each other. For instance, querying and unifying a set of graphs created by different team members enables a search for similarities, discrepancies and incoherencies.

The Scenarization module (I-SCE) deals with a transformation challenge to generate an executable model. In the selected approach, the user selects concepts and relations from the comprehension model that are relevant for a simulation. Each selected element is standardised using a predefined vocabulary while useless items are put aside. Details required for a simulation, such as properties and behaviour implementations are then added to produce an executable model.

Most simulation (executable) models are studied from simulator tailored tools with little interactivity to support their developers or users in their hypothesis testing – simulation – analysis processes. The Simulation module (I-SIM) strength resides in its very high level of interactivity enabling orderly investigation of the space of variables. It also provides a scenario-independent visual representation and interaction metaphors for supporting the user in its investigation process.

The Exploration module (I-EXP) aims at providing powerful tools allowing experts to invent views that bring a significant meaning to large datasets generated by the I-SIM module. These views have to generate, stimulate, increase or accelerate the CS understanding, and consequently augment individual or collective comprehension models. The I-EXP module puts emphasis on the proper use of technologies in order to fulfill the expert’s exploration needs. The I-EXP tools implement visualization and interaction techniques and / or mechanisms by, taking advantage of the potential of different platforms in terms of technology and in terms of exploitation of human sensory information. These tools mainly offer the user rapid prototyping of views and focus on the “best view” concept rather than the “perfect view” concept. Furthermore, to address as much as possible exploration needs with respect to diversity and richness, the toolset proposed is, if necessary, a combination of turn-key and COTS tools.

Fifty-eight people have undergone between nine (9) and fourteen hours (depending the group they were in) of effort distributed over several sessions. This experimentation was addressing three questions. Does IMAGE improve comprehension of CS? What is the added value of immersion for IMAGE? What are the typical and the efficient ways to exploit IMAGE?

Experimental results show that IMAGE facilitates the comprehension of CSs, especially when dealing with manageable level of complexity (called low level of complexity in the experiment). Observations of improvements vary according to the types of comprehension measures. More precisely, IMAGE improved the comprehension overall (not specific to any level of complexity) when the tools were available to answer questions (called analysis phase). In the case where participants had to rely only on previous knowledge acquired with the tools, IMAGE provided a benefit only at the lowest level of complexity (prediction and mental models). Additional analyses revealed that about half of the benefit of IMAGE over Baseline was attributable to the ability of IMAGE to increase comprehension; the remaining was explained by the error / mistake reduction potential of IMAGE. Experimental results do not support the added-value of immersion but this finding is not surprising considering that the experiment started before the immersive version of IMAGE was adequately implemented. In the vast majority of measures taken, participants of IMAGE group outperformed those in Baseline group on the average but many of these were not statically significant due to individual variability in this kind of experiment. This issue will have to be addressed in future work.

In addition to actual results, many lessons were learned from such an experimental context. Almost all of them are directly related to the objective of measuring the augmentation of comprehension combined to the exploratory nature of the project. This has required the development and adoption of methods that were hard to implement. Without a doubt, the IMAGE experimentation was extremely useful in terms of results but also in terms of learning for all team members. However, a second study of this sort would most likely not be as beneficial. Knowing the required time and effort to conduct such experimentation, less ambitious objectives should be considered. For instance, a case-study approach instead of a group comparison approach may be more appropriate.

Many CF contexts could potentially benefit from such a tool accelerating understanding. A whole-of-government team working out a situation involving security, development and political issues is an obvious one but facing a cyber-threat is another one requiring an efficient collaboration between experts of friendly plans, enemy intents, computer networks and software applications. The first one is operational while the second one has more a technological flavour, although it may have important operational impacts.

CF recognises the importance of a comprehensive approach considering all dimensions of a situation, more than the military power alone. The various issues (ethnics, religious, ideological and material), the various power and influence (diplomatic, economic, informational ...) as well as the national and international public opinion and Medias are examples of the variety and diversity of the dimensions. A comprehensive approach is required to elaborate plans but moreover, a whole-of-government approach is required to achieve the strategic national objectives.

Cyber is a quite new and complex challenge. Actually, CF created the new environment Cyber, distinct from Land, Air and Sea, to face it. Risk analysis identifying critical components, investigating attack vectors and study impacts of successful attacks is needed to face cyber threats impacting freedom of CF manoeuvre. Such analysis requires an efficient collaboration between experts of friendly plans, enemy intents, computer networks and software applications. Using graphs to represent potential interactions between these elements and simulation data to explore effects of attacks on the C4ISR systems would enhance CF capabilities to assess and mitigate cyber risks.

This page intentionally left blank.

References

Aissaoui, G., Genest, D., & Loiseau, S. (2003). Cognitive map of conceptual graphs: A graphical model to help for decision. *Conceptual Structures for Knowledge Creation and Communication*, 337–350.

Apache Foundation. (2010). The Apache Velocity Project. Retrieved from http://velocity.apache.org/

Argyris, C., & Schön, D. (1978). *Organizational Learning: A Theory of Action Approach* (Addison- Wesley.). Reading, MA, USA.

Arms, L., Cook, D., & Cruz-Neira, C. (1999). The benefits of statistical visualization in an immersive environment. In *Proceedings of Virtual Reality 1999* (pp. 88–95).

Ashby, W. R. (1956). *An introduction to cybernetics* (Vol. 80). Taylor & Francis.

Ayvaz, U., Dere, M., & Tiah, Y. M. (2007). Using the MANA agent-based simulation tool to evaluate and compare the effectiveness of ground-based and airborne communications jammers in countering the IED threat to ground convoys (pp. 113–118). San Diego, CA, USA: Society for Computer Simulation International.

Bailly, G. (2009). *Techniques de menus: Caractérisation, conception et évaluation*. Université Joseph Fourier - Grenoble 1.

Bazan, J., Peters, J., & Skowron, A. (2005). Behavioral pattern identification through rough set modelling. *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing*, *1*(72), 688–697.

Becker, R. A., & Cleveland, W. S. (1987). Brushing Scatterplots. *Technometrics*, *29*(2), 127–142. doi:10.1080/00401706.1987.10488204

Bell, P. C., & O'Keefe, R. M. (1987). Visual Interactive Simulation — History, recent developments, and major issues. *SIMULATION*, *49*(3), 109–116.

Bernier, F. (2009). *Comprendre des situations complexes liées aux opérations militaires*. Presented at the Matinées Sciences & Technologie RDDC Valcartier, Québec, Canada.

Bernier, F., Laurendeau, D., Rioux, F., DeRainville, F. M., & Lizotte, M. (2010). IMAGE – Interactive Simulation to Increase Complex System Understanding. In *Proceedings of the 2010 NATO Modelling and Simulation Group Workshop*. Ottawa, Canada.

Bernier, F., & Rioux, F. (2010). *Convoy scenario for complexity study: A coevolutionary perspective on convoy against insurgents* (Technical memorandum No. TM 2010-150) (p. 62). Québec, Canada: DRDC Valcartier.

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.

Brandstein, A., & Horne, G. (1998). Data farming: a meta-technique for research in the 21st century. *Maneuver Warfare Science 1998*, 93–99.

Brodlie, K., Poon, A., Wright, H., Brankin, L., Banecki, G., & Gay, A. (1993). GRASPARC-a problem solving environment integrating computation and visualization. In *Proceedings of Visualization'93* (pp. 102–109). San Jose, CA, USA.

Bryson, S. (1996). Virtual reality in scientific visualization. *Commun. ACM*, *39*(5), 62–71. doi:10.1145/229459.229467

Burdea, G. C., & Coiffet, P. (2003). *Virtual Reality Technology*. John Wiley & Sons-IEEE Press; 2 edition.

Buxton, B. (2007). Multi-touch systems that i have known and loved. *Microsoft Research*.

Cerney, M. M. (2005). *From gesture recognition to functional motion analysis: quantitative techniques for the application and evaluation of human motion*. Iowa State University.

Chein, M., & Mugnier, M. L. (2004). Concept types and coreference in simple conceptual graphs. *Conceptual Structures at Work*, 238–238.

Chein, M., & Mugnier, M. L. (2009). *Graph-based knowledge representation: computational foundations of conceptual graphs*. Springer-Verlag New York Inc.

Cioppa, T. M. (2002). *Efficient nearly orthogonal and space-filling experimental designs for high-dimensional complex models*. Monterey, CA, USA: Naval Postgraduate School.

Corbett, D. (2003). *Reasoning and unification over conceptual graphs*. Springer.

Creagh, H. (2003). Cave Automatic Virtual Environment. In *Electrical Insulation Conference and Electrical Manufacturing Coil Winding Technology Conference, 2003. Proceedings* (pp. 499 – 504). doi:10.1109/EICEMC.2003.1247937

Cressie, N. (1990). The origins of kriging. *Mathematical Geology*, *22*(3), 239–252.

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

Curtis, N. J., & Dortmans, P. J. (2004). A Dynamic Conceptual Model to Explore Technology-Based Perturbations to a Complex System: The Land Force. *Asia - Pacific Journal of Operational Research*, *21*(4), 561–563.

Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology*, *41*(1), 10–32. doi:10.1111/j.1467-8535.2009.01038.x

Davis, P. K. (2000). Dealing with complexity: exploratory analysis enabled by multiresolultion, multiperspective modeling. In *Proceedings of the Winter Simulation Conference* (pp. 293–302). Orlando, FL, USA.

Davis, P. K., Bankes, S. C., & Egner, M. (2007). *Enhancing strategic planning with massive scenario generation: Theory and experiments* (Vol. 392). RAND Corporation.

De Jong, E. (2004). Intransitivity in coevolution (pp. 843–851).

De Rainville, F.-M., Gagné, C., Teytaud, O., & Laurendeau, D. (2009). Optimizing low-discrepancy sequences with an evolutionary algorithm. In *Proceedings of GECCO'09* (pp. 1491–1498). New York, NY, USA.

DeGregory, K. W. (2007). *Optimization-based allocation of force protection resources in an asymmetric environment* (Thesis). Massachusetts Institute of Technology, Boston, MA, USA.

Dickerson, J. A., & Kosko, B. (1994). Adaptive fuzzy cognitive maps in virtual worlds. In *Proceedings of the World-Congress-on-Neural-Networkseting*. San Diego, CA, USA.

Dieng, R., & Hug, S. (1998). Comparison of personal ontologies represented through conceptual graphs. In *Proceedings of the 13th European Conference on Artificial Intelligence (ECAI 98)* (pp. 341–345). Brighton, UK.

Drolet, F. (2008). *Cohérence et synchronisation dans un environnement virtuel multi-sensoriel réparti*. Université Laval, Québec, Canada.

Edmonds, B. (2005). In *Simulation and Complexity-how They Can Relate* (Lit Verlag Munster., pp. 5–28). Münster, Germany.

Fang, K.-T., Li, R., & Sudjianto, A. (2005). *Design and Modeling for Computer Experiments*. Chapman and Hall/CRC.

Fialli, J., & Vajjhala, S. (2005). *The Java Architecture for XML Binding (JAXB) 2.0*. Sun Microsystems.

Fischer, M., Laroque, C., Huber, D., Krokowski, J., Mueck, B., Kortenjan, M., … Dangelmaier, W. (2007). Interactive refinement of a material flow simulation model by comparing multiple simulation runs in one 3d environment (pp. 499–505).

Gagnon, J. F., Jeuniaux, P., Lafond, D., & Tremblay, S. (2012). *IMAGE Cognitive System for Complexity Discovery: IMAGE V1 Experimentation Report* (Contrat report No. CR 2012-) (p. 164). Québec, Canada: RDDC Valcartier.

Genest, D. (2011). Cogitant Reference Manual version 5.2.9. Retrieved from http://cogitant.sourceforge.net/files/cogitant.pdf

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

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

Golfarelli, M., Rizzi, S., & Proli, A. (2006). Designing what-if analysis: towards a methodology. In *Proceedings of the 9th ACM International Workshop on Data Warehousing and OLAP* (pp. 51–58). New York, NY, USA: ACM.

Gore, R., Reynolds Jr., P. F., Tang, L., & Brogan, D. C. (2007). Explanation Exploration: Exploring Emergent Behavior. In *Proceedings of the 21st International Workshop on Principles of Advanced and Distributed Simulation* (pp. 113–122). Washington, DC, USA.

Government of the Islamic Republic of Afghanistan. (2006). *Afghanistan National Development Strategy - An Interim Strategy for Security, Governance, Economic Growth & Poverty Reduction* (p. 234).

Guiard, Y. (1987). Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *Journal of motor behavior*, *19*(4), 486–517.

Gunzburer, M., & Burkdart, J. (2004). *Uniformity measures for point samples in hypercubes*. Tallahassee, FL, USA: Florida State University.

Hansen, N., & Ostermeier, A. (2001). Completely derandomized self-adaptation in evolution strategies. *Evolutionary computation*, *9*(2), 159–195.

Heumer, G., Amor, H. B., Weber, M., & Jung, B. (2007). Grasp recognition with uncalibrated data gloves-A comparison of classification methods. In *Proceedings of Virtual Reality Conference* (pp. 19–26). Charlotte, NC, USA.

Hilliard, R. (2000). IEEE-Std-1471-2000 Recommended Practice for Architectural Description of Software-Intensive Systems. *IEEE, http://standards. ieee. org*.

Hinckley, K., Pausch, R., Proffitt, D., & Kassell, N. F. (1998). Two-handed virtual manipulation. *ACM Trans. Comput.-Hum. Interact.*, *5*(3), 260–302.

Holland, J. H. (1992). Complex adaptive systems. *Daedalus*, *121*(1), 17–30.

Horne, G. E., & Meyer, T. E. (2004). Data Farming: Discovering Surprise. In *Proceedings of the 2004 Winter Simulation Conference*. Washington, DC,USA.

Huot, S., & Lecolinet, E. (2007). ArchMenu et ThumbMenu: contrôler son dispositif mobile «sur le pouce». In *Proceedings of the 19th International Conference of the Association Francophone d'Interaction Homme-Machine* (pp. 107–110). Paris, France.

Hurrion, R. D. (1976). *The design, use and required facilities of an interactive visual computer simulation language to explore production planning problems*. University of London, London, UK.

Jin, Y., & Sendhoff, B. (2003). Trade-off between performance and robustness: An evolutionary multiobjective approach. In *Proceedings of the 5th International Conference* (pp. 68–68). Nantes, France.

Kent, S. (2002). Model Driven Engineering. In M. Butler, L. Petre, & K. Sere (Eds.), *Integrated Formal Methods* (Vol. 2335, pp. 286–298). Springer Berlin / Heidelberg.

Kerren, A., Ebert, A., & Meyer, J. (2007). *Human-centered visualization environments*. Springer-Verlag.

Kirkwood, C. W. (1998). System Dynamics Methods. *College of Business Arizona State University USA*.

Kleijnen, J. P. C. (1992). Regression metamodels for simulation with common random numbers: comparison of validation tests and confidence intervals. *Management Science*, *38*(8), 1164–1185.

Kleijnen, J. P. C., Sanchez, S. M., Lucas, T. W., & Cioppa, T. M. (2005). A user's guide to the brave new world of designing simulation experiments. *INFORMS Journal on Computing*, *17*(3), 263–289.

Knight, C., & Munro, M. (1999). Comprehension with[in] virtual environment visualisations. In *Program Comprehension, 1999. Proceedings. Seventh International Workshop On* (pp. 4 –11). doi:10.1109/WPC.1999.777733

Law, A. M. (2007). *Simulation modeling and analysis* (4th Edition., Vol. 3). New York, NY, USA: McGraw-Hill.

Lévesque, J.-C. (2011). *Interactions tridimensionnelles avec gants de données*. Université Laval, Québec, Canada.

Lévesque, J.-C., Laurendeau, D., & Mokhtari, M. (2011). Bimanual gestural interface for virtual environments. In *Virtual Reality Conference (VR), 2011 IEEE* (pp. 223 –224). doi:10.1109/VR.2011.5759479

Lévesque, J.-C., Laurendeau, D., & Mokhtari, M. (2013). An Asymmetric Bimanual Gestural Interface for Immersive Virtual Environments (p. 10). Presented at the 15th International Conference on Human-Computer Interaction, Las Vegas, Nevada, USA: Springer.

Lewis, M., & Sycara, K. (1993). Modeling Multispecialist Decision Making. *Advances in Human Factors Ergonomics*, *19*, 481–481.

Lizotte, M., & Moulin, B. (1989). SAIRVO: A planning System which Implement the Actem Concept. *Knowledge-Based Systems Journal*, *2*(4), 210–218.

Lizotte, M., & Rioux, F. (2010). Image-Scenarization: A computer-aided approach for agent-based analysis and design. In *Proceedings of Winter Simulation Conference (WSC)* (pp. 837–848). Baltimore, MD, USA.

Luke, S., Cioffi-Revilla, C., Panait, L., Sullivan, K., & Balan, G. (2005). MASON: A Multiagent Simulation Environment. *SIMULATION*, *81*(7), 517–527. doi:10.1177/0037549705058073

Macal, C. M., & North, M. J. (2010). Tutorial on agent-based modelling and simulation. *Journal of Simulation*, *4*(3), 151–162.

McCormic, B., Defanti, T., & Brown, M. (1987). Visualization in scientific computing-a synopsis. *IEEE Computer Graphics and Application*, *7*(7), 61–70.

McDonald, M. L., & Upton, S. C. (2005). Investigating the dynamics of competition: coevolving red and blue simulation parameters. In *Proceedings of Winter Simulation Conference* (pp. 1008–1012). Orlando, FL, USA.

Milrad, M. (2002). Using construction kits, modeling tools and system dynamics simulations to support collaborative discovery learning. *Educational Technology & Society*, *5*(4), 76–87.

Mohammed, S., & Dumville, B. C. (2001). Team mental models in a team knowledge framework: expanding theory and measurement across disciplinary boundaries. *Journal of Organizational Behavior*, *22*(2), 89–106.

Mokhtari, M., Bernier, F., Boivin, E., DuCharme, M. B., Lizotte, M., Pestov, I., & Poussart, D. (2007). *Selection Criteria for the Complex Scenarios Used in the IMAGE Project* (Technical memorandum No. TM 2007-561). Québec, Canada: DRDC Valcartier.

Mokhtari, M., & Boivin, E. (2013a). *IMAGE V1 Exploration Module* (Technical memorandum). Québec, Canada: DRDC Valcartier.

Mokhtari, M., & Boivin, E. (2013b). *IMAGE V2 Exploration Module – A Proof of Concept* (Technical memorandum). Québec, Canada: DRDC Valcartier.

Mokhtari, M., Boivin, E., & Laurendeau, D. (2013). Making Sense of Large Datasets in the Context of Complex Situation Understanding (p. 10). Presented at the 15th International Conference on Human-Computer Interaction, Las Vegas, Nevada, USA: Springer.

Mokhtari, M., Boivin, E., Laurendeau, D., Comtois, S., Ouellet, D., Levesque, J., & Ouellet, E. (2011). IMAGE - Complex situation understanding: An immersive concept development. In *Virtual Reality Conference (VR), 2011 IEEE* (pp. 229 –230). doi:10.1109/VR.2011.5759482

Mokhtari, M., Boivin, E., Laurendeau, D., & Girardin, M. (2010). Visual tools for dynamic analysis of complex situations. In *Visual Analytics Science and Technology (VAST), 2010 IEEE Symposium On* (pp. 241 –242). doi:10.1109/VAST.2010.5654451

Montgomery, D. C. (2008). *Design and analysis of experiments*. New York, NY, USA: John Wiley & Sons Inc.

Murty, K. G. (1983). *Linear programming*. New York, NY, USA: John Wiley & Sons.

Nielsen, M., Störring, M., Moeslund, T., & Granum, E. (2004). A Procedure for Developing Intuitive and Ergonomic Gesture Interfaces for HCI. In A. Camurri & G. Volpe (Eds.), *Gesture-Based Communication in Human-Computer Interaction* (Vol. 2915, pp. 105–106). Springer Berlin / Heidelberg.

Nilsson, N. J. (1980). *Principles of artificial intelligence*. Springer Verlag.

North, M. J., Howe, T. R., Collier, N. T., & Vos, R. J. (2005). The Repast Simphony Development Environment (pp. 151–158). Presented at the Agent 2005 Conference on Generative Social Processes, Models, and Mechanisms, Argonne, IL USA.

O'Connor, D. L., & Johnson, T. E. (2006). Understanding Team Cognition in Performance Improvement Teams: A Meta-Analysis of Change in Shared Mental Models. In *Proceedings of the Second Int. Conference on Concept Mapping*. San Jose, Costa Rica.

O'Connor, D. L., Johnson, T. E., & Khalil, M. K. (2004). Measuring team cognition: Concept mapping elicitation as a means of constructing team shared mental models in an applied setting. In *Proceedings of the First International Conference on Concept Mapping*. Pamplona, Spain.

Ouellet, E. (2012). *Représentation et manipulation de données de simulation dans un environnement virtuel immersif*. Université Laval, Québec, Canada.

Owen, R., Kurtenbach, G., Fitzmaurice, G., Baudel, T., & Buxton, B. (2005). When it gets more difficult, use both hands: exploring bimanual curve manipulation. In *Proceedings of Graphics Interface 2005* (pp. 17–24). School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada: Canadian Human-Computer Communications Society. Retrieved from http://dl.acm.org/citation.cfm?id=1089508.1089512

Parker, M. (2012). Agent Modeling Platform. Retrieved from http://www.eclipse.org/amp/

Philippakis, A. S. (1988). Structured what if analysis in DSS models. In *Proceedings of the Twenty-First Annual Hawaii International Conference on Decision Support and Knowledge Based Systems Track* (Vol. 3, pp. 366–370). Kailua-Kona, HI, United States.

Pickles, S., Haines, R., Pinning, R., & Porter, A. (2004). Practical tools for computational steering. In *Proceedings of the UK e-Science All Hands Meeting* (Vol. 31). Nottingham, UK.

Polovina, S. (1993). *The Suitability of Conceptual Graphs in Strategic Management Accountancy*. Loughborough University of Technology, Leicestershire, UK.

Raja, D., Bowman, D., Lucas, J., & North, C. (2004). Exploring the benefits of immersion in abstract information visualization. In *Proceedings of the Immersive Projection Technology Workshop*. Ames , IA, USA.

Ribière, M., Matta, N., & Cointe, C. (1998). A proposition for managing project memory in concurrent engineering. *Interpretation*, *2*(1), 2.

Richardson, K. A., Mathieson, G., & Cilliers, P. (2000). The theory and practice of complexity science: Epistemological considerations for military operational analysis. *SysteMexico*, *1*(1), 25–66.

Rioux, F. (2008a). *Conception et mise en oeuvre de multichronia, un cadre conceptuel de simulation visuelle interactive*. Université Laval, Québec, Canada.

Rioux, F. (2008b). Visual Interactive Simulation and Data Farming : How Can They Co- Exist ? In *Proceedings of the International Data Farming Workshop*. Monterey, CA, USA.

Rioux, F., Bernier, F., & Laurendeau, D. (2008a). Multichronia–A Generic Visual Interactive Simulation Exploration Framework. In *Procedings of The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC)*. Orlando, FL, USA.

Rioux, F., Bernier, F., & Laurendeau, D. (2008b). Visualising and interacting with multiple simulations using the multichronic tree. In *Proceedings of the International Conference on Modeling, Simulation and Visualization Methods*. Las Vegas, NV, USA.

Rioux, F., Bernier, F., & Laurendeau, D. (2008c). Design and implementation of an XML-based, technology-unified data pipeline for interactive simulation. In *Proceedings of Winter Simulation Conference (WSC)* (pp. 1130–1138). Miami, FL, USA.

Rioux, F., Laurendeau, D., & Bernier, F. (2010). Visualising and interacting with multiple simulations using the multichronic tree. *International Journal of Computer Aided Engineering and Technology*, *2*(1), 52–65.

Rioux, F., & Lizotte, M. (2011). Image-Scenarization: From conceptual models to executable simulation. In *Proceedings of Winter Simulation Conference (WSC)* (pp. 271–283). Phoenix, AZ, USA.

Rothrock, L., & Narayanan, S. (2011). *Human-In-The-Loop Simulations: Methods and Practice*. Springer.

Sadowski, W., & Stanney, K. (2002). Presence in virtual environments. In *Handbook of Virtual Environments: Design, Implementation, and Applications* (pp. 791–806). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Santner, T. J., Williams, B. J., & Notz, W. (2003). *The design and analysis of computer experiments*. New York, NY, USA: Springer Verlag.

Schlattmann, M., & Klein, R. (2009). Efficient bimanual symmetric 3D manipulation for markerless hand-tracking. In *Proceedings of the Virtual Reality International Conference (VRIC)*. Laval, France.

Schulz, H.-J. (2011). Treevis.net: A Tree Visualization Reference. *Computer Graphics and Applications, IEEE*, *31*(6), 11 –15. doi:10.1109/MCG.2011.103

Sherman, W. R., & Craig, A. B. (2003). Understanding Virtual Reality—Interface, Application, and Design. *Presence: Teleoperators and Virtual Environments*, *12*(4), 441–442. doi:10.1162/105474603322391668

Smola, A., & Vapnik, V. (1997). Support vector regression machines. *Advances in neural information processing systems*, *9*, 155–161.

Sowa, J. F. (1998). *Conceptual Graph Standard and Extensions*. In *Proceedings of the 6th International Conference on Conceptual Structures*, Montpellier, France.

Sowa, J. F. (1984). *Conceptual Structures: Information Processing in Mind and Machine*. Reading: Addison-Wesley.

Spall, J. C. (1998). An overview of the simultaneous perturbation method for efficient optimization. *Johns Hopkins APL Technical Digest*, *19*(4), 482–492.

Standridge, C. R., Matwiczak, K. M., Davis, D. A., Musselman, K. J., & Brunner, D. T. (1990). Interactive simulation. In *Proceedings of the Winter Simulation Conference* (pp. 453–458). New Orleans, Louisiana, USA.

Sterman, J. D. (2000). *Business dynamics: systems thinking and modeling for a complex world* (Vol. 53). Irwin McGraw-Hill.

Storey, M. A. (2006). Theories, tools and research methods in program comprehension: past, present and future. *Software Quality Journal*, *14*(3), 187–208.

Tye-Gingras, O. (2011). *Intégration et exploitation d'outils de visualisation - Compréhension de situations complexes*. Université Laval, Québec, Canada.

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

Veit, M., Capobianco, A., & Bechmann, D. (2008). Consequence of Two-handed Manipulation on Speed, Precision and Perception on Spatial Input Task in 3D Modelling Applications. *Journal of Universal Computer Science*, *14*(19), 3174–3187.

Wickens, C. D., Gordon, S. E., & Liu, Y. (2004). *An introduction to human factors engineering*. Pearson Prentice Hall.

Wright, H. (2004). Putting visualization first in computational steering. In *Proceedings of UK e-Science All Hands Meeting* (pp. 326–331). Nottingham, UK.

Youngblut, C. (2006). *What a Decade of Experiments Reveals about Factors that Influence the Sense of Presence*. Retrieved from http://stinet.dtic.mil/oai/oai?&verb=getRecord&metadataPrefix=html&identifier=ADA473402

This page intentionally left blank.

Annex A Questionnaire Flash (in french)

A.1 Questionnaire tactique

Quelle attribution de ressources :

a. Permet la *meilleure* **ponctualité** moyenne tout en conservant une **intégrité** du convoi *d'au moins 8.5* ?

b. Génère le plus *faible* **cargo** du convoi ?

c. Permet d'avoir une **ponctualité** *d'au moins 6.5* étant donné une **charge utile** *maximale de 0.8* et un **cargo** *d'au moins 7.2* ?

d. Génère les plus *faibles* **mesures de performance** du convoi en moyenne ?

Réponses attendues pour chaque question

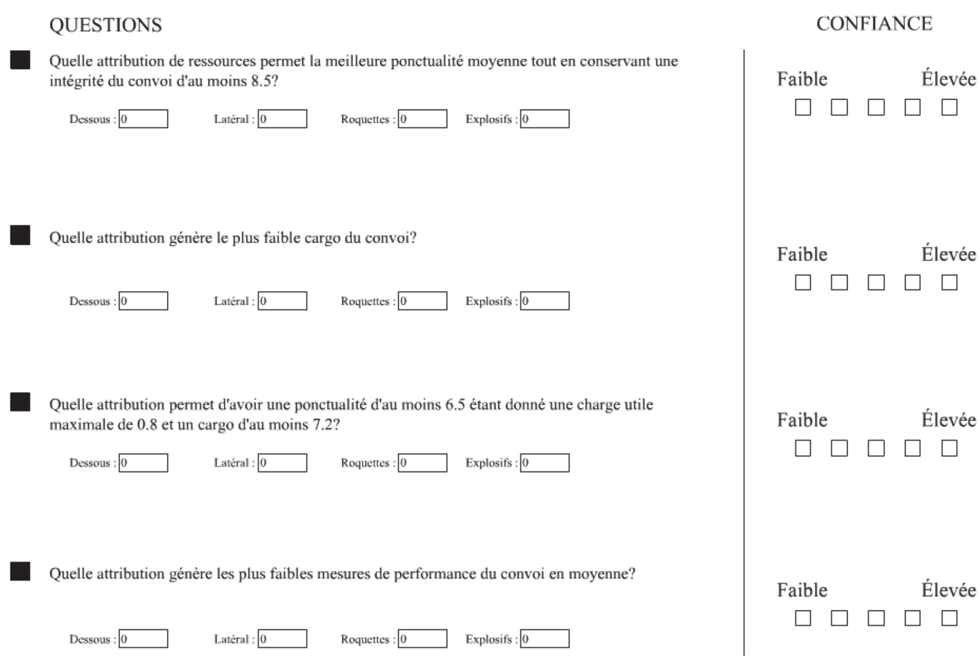
i. Blindage dessous

ii. Blindage latéral

iii. Nombre roquettes

iv. Nombre explosifs

v. Niveau de confiance globale



QUESTIONS — CONFIANCE

Quelle attribution de ressources permet la meilleure ponctualité moyenne tout en conservant une intégrité du convoi d'au moins 8.5?

Dessous : 0 Latéral : 0 Roquettes : 0 Explosifs : 0 — Faible ☐ ☐ ☐ ☐ ☐ Élevée

Quelle attribution génère le plus faible cargo du convoi?

Dessous : 0 Latéral : 0 Roquettes : 0 Explosifs : 0 — Faible ☐ ☐ ☐ ☐ ☐ Élevée

Quelle attribution permet d'avoir une ponctualité d'au moins 6.5 étant donné une charge utile maximale de 0.8 et un cargo d'au moins 7.2?

Dessous : 0 Latéral : 0 Roquettes : 0 Explosifs : 0 — Faible ☐ ☐ ☐ ☐ ☐ Élevée

Quelle attribution génère les plus faibles mesures de performance du convoi en moyenne?

Dessous : 0 Latéral : 0 Roquettes : 0 Explosifs : 0 — Faible ☐ ☐ ☐ ☐ ☐ Élevée

A.2 Questionnaire procédural

Quelle est la meilleure série de choix dans le but de maximiser :

a. L'intégrité totale ?

b. Le cargo total ?

c. La ponctualité totale ?

Réponses attendues pour chaque question

i. Allocation initiale

Blindage dessous, Blindage latéral, % de don, Nb explosifs Nb roquette, Priorité

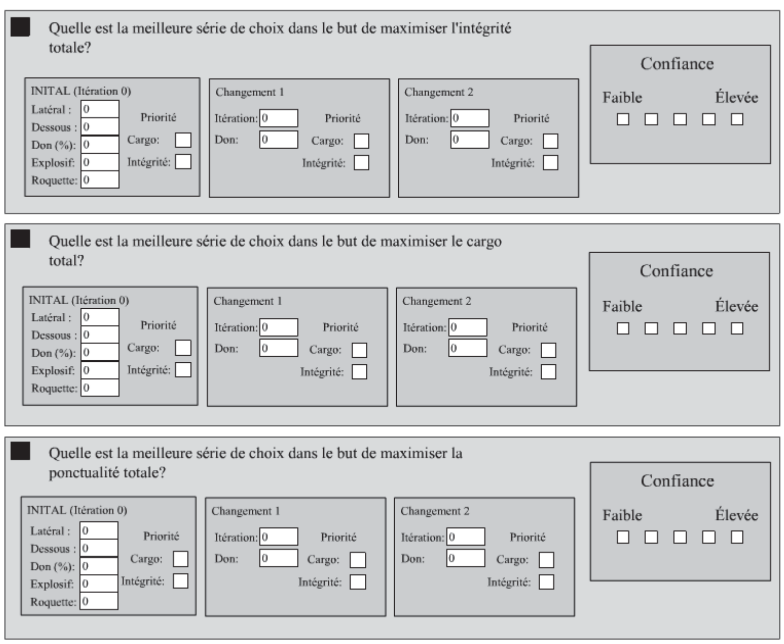
ii. Changement 1

Itération, % de don, Priorité

iii. Changement 2

Itération, % de don, Priorité

iv. Niveau de confiance globale


Quelle est la meilleure série de choix dans le but de maximiser l'intégrité totale?

INITAL (Itération 0) — Latéral : 0; Dessous : 0; Don (%): 0; Explosif: 0; Roquette: 0; Priorité: Cargo: ☐ Intégrité: ☐

Changement 1 — Itération: 0; Don: 0; Priorité: Cargo: ☐ Intégrité: ☐

Changement 2 — Itération: 0; Don: 0; Priorité: Cargo: ☐ Intégrité: ☐

Confiance — Faible ☐ ☐ ☐ ☐ ☐ Élevée

Quelle est la meilleure série de choix dans le but de maximiser le cargo total?

INITAL (Itération 0) — Latéral : 0; Dessous : 0; Don (%): 0; Explosif: 0; Roquette: 0; Priorité: Cargo: ☐ Intégrité: ☐

Changement 1 — Itération: 0; Don: 0; Priorité: Cargo: ☐ Intégrité: ☐

Changement 2 — Itération: 0; Don: 0; Priorité: Cargo: ☐ Intégrité: ☐

Confiance — Faible ☐ ☐ ☐ ☐ ☐ Élevée

Quelle est la meilleure série de choix dans le but de maximiser la ponctualité totale?

INITAL (Itération 0) — Latéral : 0; Dessous : 0; Don (%): 0; Explosif: 0; Roquette: 0; Priorité: Cargo: ☐ Intégrité: ☐

Changement 1 — Itération: 0; Don: 0; Priorité: Cargo: ☐ Intégrité: ☐

Changement 2 — Itération: 0; Don: 0; Priorité: Cargo: ☐ Intégrité: ☐

Confiance — Faible ☐ ☐ ☐ ☐ ☐ Élevée

List of symbols/abbreviations/acronyms/initialisms

ACE	Adaptive communication environment
ANOVA	Analysis of variance
ANSI	American National Standards Institute
AVG	Average
BF	Blue forces
CAVE	CAVE automatic virtual environment
CF	Canadian Forces
CMA-ES	Covariance matrix adaptation evolution strategy
COTS	Commercial off-the-shelf
CS	Complex situation
DND	Department of National Defence
DOF	Degree of freedom
DRDC	Defence Research & Development Canada
DRDKIM	Director Research and Development Knowledge and Information Management
GB	Gigabytes
HD	Hard drive
GHz	Gigahertz
I-EXP	IMAGE Exploration
I-REP	IMAGE Representation
I-SCE	IMAGE Scenarisation
I-SIM	IMAGE Simulation
IDoE	Interactive design of experiment
IED	Improvised explosive device
IEEE	Institute of Electrical and Electronics Engineers
IVE	Interactive Virtual Environment
JAXB	Java architecture for XML binding
LAV	Light armoured vehicle
LCD	Liquid crystal display
MoP	Measure of performance

NIED	Number of improvised explosive devices
NRPG	Number of rocket-propelled grenades
OSG	OpenSceneGraph
PC	Personal computer
R&D	Research & Development
RAM	Random access memory
RF	Red forces
RPG	Rocket-propelled grenades
RSF	Response selection fallacy
SAT	Side armour vehicle
SDK	Software development kit
SOP	Standard operating procedure
SPSA	Simultaneous perturbation stochastic approximation
SQL	Standard query language
SVM	Support vector machine
VIS	Visual interactive simulation
VR	Virtual reality
XML	Extensible Markup Language
XSLT	Extensible Stylesheet Language Transformations

DOCUMENT CONTROL DATA (Security markings for the title, abstract and indexing annotation must be entered when the document is Classified or Designated)	
1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's report, or tasking agency, are entered in section 8.) Defence Research and Development Canada – Valcartier 2459 Pie-XI Blvd North Quebec (Quebec) G3J 1X5 Canada	2a. SECURITY MARKING (Overall security marking of the document including special supplemental markings if applicable.) UNCLASSIFIED 2b. CONTROLLED GOODS (NON-CONTROLLED GOODS) DMC A REVIEW: GCEC JUNE 2010

3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title.) IMAGE Final Report : An Interactive Computer-Aided Cognition Capability		
4. AUTHORS (last name, followed by initials – ranks, titles, etc. not to be used) Lizotte, M.; Bernier, F.; Mokhtari, M.; Boivin, É.		
5. DATE OF PUBLICATION (Month and year of publication of document.) December 2012	6a. NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.) 140	6b. NO. OF REFS (Total cited in document.) 126
7. DESCRIPTIVE NOTES (The category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.) Technical Report		
8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.) Defence Research and Development Canada – Valcartier 2459 Pie-XI Blvd North Quebec (Quebec) G3J 1X5 Canada		

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.) 10bj	9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)
10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.) DRDC Valcartier TR 2012-397	10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)
11. DOCUMENT AVAILABILITY (Any limitations on further dissemination of the document, other than those imposed by security classification.) Unlimited	
12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11) is possible, a wider announcement audience may be selected.)) Unlimited	

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

With the advances in Science and Technology (S&T), we are now able to build artefacts whose complexity approaches those of life itself. To address the new defence and national security challenges of dealing with an ever increasing range of possibilities – some natural but many human-made – DRDC scientists propose a toolset concept that enables the collaboration of experts from different disciplines so that scientists can try to reach a shared understanding of a complex situation. The toolset concept supports the need for disentangling complex situations more quickly by using the synergy between knowledge representation, modeling, simulation, and visualisation technologies. Many Canadian Forces contexts could potentially benefit from such a toolset in order to accelerate understanding. A whole-of-government team working out a situation involving security, development, and political issues is an obvious context, but facing a cyber-threat requires an efficient collaboration between experts in order to analyze friendly plans, enemy intents, computer networks, and software applications.

Avec les progrès en sciences et technologie (S&T), l'homme est désormais capable de construire des objets dont la complexité se rapproche de celle de la vie elle-même. Pour aborder les nouveaux défis de sécurité nationale et de défense devant traiter un éventail toujours croissant d'éléments, certains naturels, mais beaucoup fabriqués par l'homme, les scientifiques de RDDC proposent un concept de boîte à outils visant la collaboration d'experts œuvrant dans différentes disciplines qui tentent de parvenir à une compréhension commune d'une situation complexe. Ce concept vise à soutenir le besoin de déchiffrer des situations complexes plus rapidement en utilisant la synergie entre la représentation des connaissances, la modélisation, la simulation et les technologies de visualisation. Les Forces canadiennes pourraient potentiellement bénéficier de ces outils pour accélérer la compréhension de nombreux contextes dans lesquels elles doivent être déployées. Une équipe pangouvernementale travaillant sur une situation impliquant des questions politiques, de sécurité et de développement est un contexte évident. Par contre, faire face à une menace cybernétique est un contexte nécessitant une collaboration efficace entre les experts de la planification côté ami, les experts des intentions de l'ennemi, ainsi que ceux des réseaux informatiques et des applications logicielles.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

complexity; comprehension; cognition; military; modeling; simulation; visuatlisation; representation; virtual reality

Defence R&D Canada

Canada's Leader in Defence
and National Security
Science and Technology

R & D pour la défense Canada

Chef de file au Canada en matière
de science et de technologie pour
la défense et la sécurité nationale

DEFENCE DÉFENSE

www.drdc-rddc.gc.ca