Chapter 12 – TECHNICAL EVALUATION REPORT

Mr. William Fraser
Human Factors Engineering Group
Human Systems Integration Section
Defence Research and Development Canada – Toronto
P.O. Box 2000, 1133 Sheppard Ave. West
Toronto, Ontario M3M 3B9
CANADA

12.1 EXECUTIVE SUMMARY

The Specialists' Meeting HFM-143/RSM on Human Behaviour Representation (HBR) in Constructive Modelling was held at the Canadian Forces College, Toronto, Canada from 30-31 May, 2007. The goal of the meeting was to obtain input from the scientific and operational communities regarding the validity of the recommendations made by the HFM-128 Task Group and to solicit input in those areas where the HFM-128 Task Group lacked experience. The meeting was organized around seven topic areas of importance in developing useful human behaviours:

1) What human factors does the operation involve;
2) Human task representation in M&S;
3) Behaviour generation – variability and choice;
4) Military relevant mental output measures;
5) The concept of moderators;
6) Complexity, hierarchy, modularity, and validity in HBR Architectures; and
7) From individual to group behaviour.

The structure of the meeting, with the focus on modelling issues rather than presentations on specific modeling tools or models, provided the forum and the time for a wide-ranging discussion of the many and complex problems and challenges faced by those involved in Human Behaviour Representation in constructive modelling and simulation. One of the central themes arising from the presentations and the discussions is the increasing need to incorporate a wide-range of Human Behavioural Representations into constructive simulations of complex, non-kinetic, operations involving a mix of military and civilian organizations. It is also apparent that there is a serious deficiency in tools for advanced model development, model sharing, simulation, and validation. To address these and other issues, NATO RTO should initiate a series of initiatives to:

1) Address the development of advanced modelling tools to expand the capabilities in simulating individuals, individuals within teams and organizations, and teams and organizations;
2) Develop new approaches to performance and behaviour metrics in large scale simulations and real-world operations;
3) Develop mechanisms to share data within the NATO community that has been collected from both small and large scale real-world operations, human-in-the-loop simulations, and constructive simulations, to aid in model development and validation; and

4) Develop technologies for large scale automated data reduction, analysis, and display, utilizing advances in GRID technologies.

12.2 INTRODUCTION

The Specialists' Meeting HFM-143/RSM on Human Behaviour Representation (HBR) in Constructive Modelling was held at the Canadian Forces College, Toronto, Canada from 30-31 May, 2007. The panel was organized by the Human Factors and Medicine Panel in coordination with the members of the HFM-128/ RTG.

Constructive simulation is becoming an ever more popular and powerful tool for use in Modelling and Simulation environments within NATO community. The representation of human behaviour and performance in these constructive simulations is a critical factor in the use of the technology to enhance the capability of the overall M&S environment. However, a large number of difficult and challenging issues surround the development and implementation of these representations. The purpose of the HFM-128 Task Group is to make recommendations on the implications of human factors and cognitive science to improve the implementation of human behaviour representations in complex simulations. The goal of the HFM-143 Specialists' Meeting was to obtain input from the scientific and operational communities regarding the validity of the recommendations made by the HFM-128 Task Group and to solicit input in those areas where the HFM-128 Task Group lacked experience.

In order to simulate discussions and to provide focused feedback to the HFM-128 Task Group, the meeting was organized around seven topic areas of importance:

1) What human factors does the operation involve;
2) Human task representation in M&S;
3) Behaviour generation – variability and choice;
4) Military relevant mental output measures;
5) The concept of moderators;
6) Complexity, hierarchy, modularity, and validity in HBR Architectures; and
7) From individual to group behaviour.

Formal presentations were relatively few and short. Instead, introductory comments by the Chair of each session were designed to stimulate discussion and feedback on the topics of concern. This technical review has attempted to focus on some general themes that emerged during the two days of discussion which suggest future areas of research focus for NATO RTO. Several specific topics for follow on Tasking Groups are identified.

12.3 SUMMARY

The first day of the meeting started with an introduction by Dr. Lotens, Chair of the HFM-128 Task Group and a Keynote address by Dr. Foster, Director, BioSystems, Office of the Director, Defense Research and Engineering, US Dept of Defense. The second day of the meeting was introduced with a second Keynote address, by Mike Greenley, Vice President of Modelling and Simulation, CAE.

Dr. Lotens' opening remarks provided a clear focus for the workshop, outlining the goals and the desired output to assist the panel in making recommendations to NATO concerning simulation of human behaviour in a military context. He reviewed the specific issues the workshop was hoping to address including some of the difficulties in modelling complex human behaviours:

1) Balanced human factors input to operational studies;
2) Generation of likely behaviours and selection of behaviour;
3) How to tailor models fit for purpose;
4) Performance metrics;
5) Reduction of complexity;
6) Reduction of effort; and
7) Reuse of developments.

He emphasized the need to develop the right model for the operator, at the right time, for the right cost.

Dr. Foster's keynote address focused on the challenges facing modelling and Human behavioural Representation given the broad range of scope, timelines, and complexity of tactical and strategic operations that are now undertaken by NATO countries still dealing with conventional and asymmetrical adversaries in the form of insurgency and terrorist operations, but more and more addressing issues of the non-kinetic battlespace with its complex socio-political issues, and interactions between military and civilian organizations who are dealing with the full range of security, stability, transition, and reconstruction (SSTR) issues. He also discussed the need for Human behavioural Representation to:

1) Extend beyond experimentation into areas of training, mission rehearsal, planning, and forecasting;
2) Expand the range of scenario and behaviour authorship to the user of modelling tools;
3) Incorporate the physiology of the operator rather than rely exclusively on empirically derived performance moderators;
4) Include complex social and cultural dynamics;
5) Develop better validation and verification approaches; and
6) Reduce the cost and time to create constructive simulations.

Dr. Foster identified three technologies that could assist this research and development: advances in gaming technology, supercomputing capability through cluster/grid based technology, and the development of open databases of real-world data to assist model development and validation.

12.4 SESSION SUMMARIES

12.4.1 Topic 1 – What Human Factors Does the Operation Involve?

Dr. Castor lead a discussion that addressed the problem of determining the level of fidelity, abstraction, and simplification of the real world that is appropriate for the modelling environment and for the particular model that is required. A major issue is determining what human factor issues should be incorporated into these simplified representations when dealing with the conflict between client needs and available modelling

resources. A number of other issues were discussed including the lack of real-world data for modelling and modelling validation, and the need for highly trained, experienced modellers with multiple domain knowledge.

12.4.2 Topic 2 – Human Task Representation in M&S

Dr. Allender lead off the discussion with a review of the types of task analyses that are done on real-world systems to assist in the modelling environment, and outlined some of the issues in the representation of tasks: level of granularity in describing tasks, role of context in defining tasks, defining tasks for teams and large scale organizations, and the importance of natural language processing tasks. Dr. Vartanian's presentation on unconscious, naturalistic decision making and task performance in complex, rapidly evolving situations, lead to an extended discussion on the problems of incorporating such complex tasks, and the type of unconscious decision making involved in doing these task, into architectures for Human Behaviour Representations that have traditionally focused on rule-based approaches. The issues and difficulties involved in incorporating emotion and creativity were also discussed.

12.4.3 Topic 3 – Behaviour Generation: Variability and Choice

Mr. Armstrong introduced the session by discussing some of the requirements of Human behavioural Representation, citing the need for models that are more complex, flexible, plausible, and capable of extension and learning to supported unscripted behaviours, as well as demonstrate the wide variability of individual and team behaviours. Two presentations were given by Mrs. Cooper-Chapman and Mr. Emiel on modelling approaches that incorporate some of the concepts of variability and choice. The UK Combat System Engineering Model focuses on situational awareness, where the experience of the operator is a main modulator of behaviour, and the TNO model incorporates the impact of the external state of the scenario and the internal state of the operator on generating stress, along with the specifics of the task/mission to determine overall operator behaviour. Following the presentations there was extensive discussion on the approaches to use in large scale simulations and the inherent difficulties in bounding the variability within the behavioural representations. It was emphasized that the skill and experience of the modeller is a critical factor in developing appropriate simulations.

12.4.4 Day 2 – Keynote Address

Mr. Mike Greenley's Keynote address discussed the Human behavioural Representation issues from an industry perspective. He noted several major trends in the development of M&S technologies, including more commercialization of the tools and delivery of complete turnkey simulation environments to the end-user with continuous ongoing facility and simulation support by the contractors and the move to simulations involving a broader range of military and non-military activities, including complex multi-person teams, groups, and organizations. He stressed that more and more intelligent agent technologies will be used in the computer generated forces of the large scale simulations and identified the increasing role of the gaming industry and the advances in gaming technology in constructive and human-in-the-loop simulations. There was discussion following Mr. Greenley's address on the issues of surrounding the rapid delivery of simulation environments and the role of intelligent agents in accelerating the development of training capabilities.

12.4.5 Topic 4 – Militarily Relevant Mental Output Measures

Dr. Belyavin lead an extensive discussion on the metrics of Workload (WL) and Situational Awareness (SA) focused on seven questions:

1) Is Workload a well-defined construct?
2) Do we need to model it and if-so how?
3) Do we need to validate workload predictions from models?
4) Is SA a well-defined construct?
5) Can we validate measures of SA?
6) Is that a useful activity?
7) Are there other measures?

It was emphasized that there is often poor correlation of the metrics with task-performance, as well as the often poor correlation between the metrics generated by the cognitive architectures used for system models and the metrics obtained from the operators of the actual systems. There was agreement that most model architectures/environments are not designed to produce estimates of the metrics, especially SA, as they focus on performance prediction. Dr. Allender supported the need for reporting SA and WL metrics, as it provides the client with valuable information. There was a general consensus that the WL and SA constructs cannot be discarded, but more work is needed to better define and validate the SA construct.

12.4.6 Topic 5 – The Concept of Moderators

Dr. Belyavin also led a wide-ranging discussion on the moderators of human behaviour focusing on four specific questions:

1) For what applications do we need to solve the moderator problem?
2) Can we ignore moderators across a range of problems?
3) How do we develop valid models if we need to?
4) How do we deal with potential moderator interactions?

Though the impact of external stresses such as the environment are recognized as key moderators, much more extensive research and data is needed on other moderators, especially on the issues of:

1) Moderator characterization;
2) The interaction of moderators such as experience and personality type with other moderators such as the environment; and
3) Traditional and new definitions of moderators in the context of teams, organizations, and complex military and civilian operations.

It was emphasized that there is no simple “state” description of human capability. Rather, “state” is a complex function of task, training, personality, environment, etc.

12.4.7 Topic 6 – Complexity, Hierarchy, Modularity and Validity in HBR Architectures

Dr. Gluck introduced some of the issues surrounding the design and structure of the currently used cognitive architectures such as ACT-R and SOAR, emphasizing the slow pace of development of cognitive architectures and modelling environments. Dr. Jones’ presentation on Comparative Analysis of Frameworks for Knowledge-Intensive Intelligent Agents, emphasized the lack of modularity of the architectures which would allow better

reuse of model components, the difficulty in separating theory and implementation in the architectures, the inability to transfer cognitive models between architectures, and the difficulties in adding intelligent agent capabilities to the current architectures. He emphasized the need to follow a rigorous software engineering approach in building new architectures for intelligent behaviour that address understandability and usability challenges from the outset and which would include comprehensive support for necessary knowledge representations, and well-defined software components. He also outlined a long term strategy to meet these goals, that included the development of an abstract machine based on common functional components, interoperable object libraries as instantiations of abstract components, a formal framework to provide a bridge between science and implementation, and an ability to compose components quickly into the "best" architecture for a given task.

Dr. West's presentation examined the different ways of looking at a model's validity including face, model, predictive, architectural, and cross-model comparative validity and also emphasized the need to link the models back to the neural architecture of the brain. He also raised the difficulties of publishing and sharing the models. The presentations were followed by an animated discussion on how to address the shortcomings of the current architectures. There appeared to no consensus on how to proceed, though the current restrictions on model development and sharing were recognized.

12.4.8 Topic 7 – From Individual to Group Behaviour

Mr. Cain opened the session with a talk on the complexities and difficulties in extending models of Human behaviour to include group behaviours. He proposed three problem/discussion areas:

1) Can we define those essential attributes of teams that require formal models that would make team entity modelling viable and are validated formal models available?
2) Can we model teams of individuals at sufficient resolution and validate them for use in military simulations and what can make this process affordable?
3) Can we define the pros and cons of each approach sufficiently so that the military M&S community can make an informed decision about which is more appropriate in a given context?
4) How is an organization representation different from a small team and do we need to represent explicit EBO behaviours or is it sufficient to represent their "effects"?

Modelling teams as both individuals and collectives can be appropriate, however models of team leaders or intelligent agents acting to coordinate activities is a critical modelling issue. There is active research in crowd modelling, and socio-cultural team modelling, but model approaches for complex, non-kinetic mixes of military and civilian organizations is lacking.

12.5 COMMENTS AND RECOMMENDATIONS

The structure of the meeting, with the focus on modelling issues rather than presentations on specific modelling tools or models, provided the forum for a wide-ranging discussion of the many and complex problems and challenges faced by those involved in Human behaviour Representation in constructive modelling and simulation. One of the central themes arising from both Keynote addresses is the increasing need to incorporate a wide-range of Human behavioural Representations into constructive simulations of complex, non-kinetic, operations involving a mix of military and civilian organizations (both government and NGOs) and civilian populations, taking into account broad socio-political issues. It is also apparent that while more work is required on developing individual and team models, there is a serious deficiency in tools for

advanced model development, model sharing, simulation, and validation – key factors in the development of useful simulations of large, complex systems. Modelling software technologies and database capabilities must be advanced to assist the community in rapidly developing, testing, validating, sharing, and using human behaviour and performance models that can be integrated into the full spectrum of M&S capabilities. It is also apparent that a generic component based approach to Human behavioural Representation is required that will facilitate a plug-and-play approach to implementing human models in large, complex constructive and HITL simulations involving system-of-systems. This effort will require the development of new modelling tools, as well as the embedding of HLA and other interfacing capabilities into existing technologies. NATO should focus less on human based experimentation, more on real-world data collection and the development of open-access databases, and more on development of general behavioural models that can be incorporated into a broad range of constructive simulations.

Though there is a desire to move authorship of simulations more and more to the client organizations, it must be recognized that the validity of the models are highly dependent on the training, skill, and experience of the modelers, who require domain knowledge in modelling theories and methodologies, in specialized software and programming languages, in mathematics and statistics, as well as in all aspects of the real-world situation they are attempting to simulate. When dealing with complex system-of-systems, the multi-domain knowledge requirements can be overwhelming. In spite of the advances in computing power, developing and using advanced Human behaviour Representations will remain an expensive and labor intensive discipline.

In order to enhance the development of Human behavioural Representations, and facilitate their incorporation into the larger M&S capabilities, NATO RTO should consider the establishment of several Working Groups to focus on specific raised by the Specialists' Meeting.

NATO RTO should initiate a series of initiatives to address the development of advanced modelling tools to expand the capabilities in simulating individuals, individuals within teams and organizations, and teams and organizations. This will require the integration of models of emotion and creativity and more emphasis on biological/physiological bottom-up models of brain function. One possible approach is develop human behaviour and performance models, using an open-source modelling tool such as Modelica, specifically designed for multi-domain, system-of-systems modelling, capable of handling a mix of mathematical representations, and allowing for the integrated modelling of humans, hardware, and communication components, as well as complex socio-political-economic systems This effort should also focus on developing tools for sharing models among different authoring tools by the development of XML schemas for model description and developing software to link multiple modelling technologies including Modelica, IPME, IMPRINT, SOAR, ACT-R, fuzzy-logic rule based systems, and neurophysiological based modelling environments such as PDP++/leabra++ .

In order to develop and validate performance and behavioural models, simulations and real-world experimentation there is a need to use well defined metrics, beyond the current individual Workload and Situational Awareness measures, which are difficult to define in individuals, and poorly defined in terms teams, groups, and networked operations. NATO RTO should establish Working Groups to develop new approaches to performance and behaviour metrics in large scale constructive simulations and real-world operations. Given the shear complexity of the large scale simulations of military and large scale civilian operations, this should include the development of technologies, including intelligent agents, for large scale automated data reduction, analysis, and display, utilizing advances in GRID technologies.

In parallel to the work on simulation, experiment, and real-world metrics, NATO should establish a Working Group on the issues of data collection and data storage, to handle the immense volume of data collected on

human behaviour and performance. In addition to the issues surrounding raw data storage, the Working Group should address database methodologies for storing information on data collection methodologies used, data format and data conversion techniques, calibration procedures, and model databases.