Serious Gaming for Complex Decision Making

Anja v.d. Hulst & Mark Ruijsendaal

TNO, P.O.Box 23, 3769 ZG Soesterberg, The Netherlands

anja.vanderhulst@tno.nl mark.ruijsendaal@tno.nl

Abstract.
Tactical- and strategic decision making in the safety domain is a form of 'complex decision making' with Naturalistic Decision Making as the predominant line of research. At the heart of the Decision Making expertise are 'situation assessment' capabilities, the most 'intuitive' aspect of complex decision making. In training it is also the most neglected. Particularly for developing the highly intuitive assessment skills, substantial task experience is indispensable. This makes serious gaming an attractive alternative to live training sites for tasks that are dangerous, hard or just too expensive. However, gaming requires a dedicated training approach like Job Oriented Training to be effective. We learned that implementing JOT in realistic settings, many lower level design issues emerge. Design choices are found to have substantial impact on the effects of training. Unsolved design issues are: level of fidelity, scenario progression design and designing for flow.

Keywords: Serious Gaming, Complex Decision Making, Situation Assessment, Job Oriented Training, Design Issues

1 Complex decision making

Tactical- and strategic decision making in the safety domain is highly situational, cognitively complex and is performed under demanding circumstances. As such, it is typically a form of 'complex decision making' (CDM) (Orasanu & Connolly, 1993). That is, decision making under circumstances that can be characterized by dynamic and continually changing conditions, uncertainty and ambiguity, ill-defined tasks, time constraints and most important, high stakes, multiple actors and significant personal consequences of mistakes (Klein, 2003).

From the 1960s, it became clear that a substantial part of failure in military operations resulted from inadequate situation awareness and decision making by commanders. From then on, the US Army began funding decision making research during the mid-1980s. The U.S. Navy became involved following the 1988 USS Vincennes shoot-down incident, in which a U.S. Navy Aegis cruiser destroyed an Iranian commercial airliner, mistaking it for a hostile attacker (Klein, 1998). TNO participated extensively almost right from the beginning in this research, e.g. by the work of Schraagen (1997, 2008). At present, in the field of military- and safety research, a vast body of research is emerging that aims at understanding complex decision making to better prepare commanders for the demanding circumstances of the safety domain.

The predominant line of research in complex decision making is that of Naturalistic Decision Making (NDM) (Klein, 1998, 2003, Zsambok et al., 1997). Research in NDM aims at developing models that describe how experienced decision makers actually function under demanding circumstances. The analyses done in the field of NDM reveal, for instance, that at the tactical and strategic level, very few general procedures exist. For example, Kahneman et al.(1982) observed that expert decision makers did not adhere to the principles of optimal performance; they relied on heuristics as opposed to algorithmic strategies. Also, it was observed that decision making is largely situated. That is, proper decision making is highly dependent on awareness of characteristics of the local terrain, infrastructure, population, the local risks as well as mandates and responsibilities of the actors involved.

2 Situation Assessment

Even more important, one of the essential findings in NDM is that 'situation assessment' (SA) capabilities are at the heart of the expertise. SA is the most 'intuitive' aspect of complex decision making and it takes most time and experience to develop to an expert level (Stehouwer et al., 2005). What distinguishes an expert from a novice in situation assessment is that the expert has acquired a comprehensive repertoire of patterns (Klein, Calderwood, & Clinton-Cirocco, 1986). 'These patterns describe the primary causal factors operating in the situation. The patterns highlight the most relevant cues, provide expectancies, identify plausible goals, and suggest typical types of reactions in that type of situation' (Klein, 2008, p. 457). Decision makers recognize and categorize situations on the basis of the patterns they have acquired, hence the label 'Recognition Primed Decision Making' (RPDM).

Situation assessment may be essential to expertise in CDM, sadly, in training it is also the most neglected (Stehouwer, 2005). Particularly for developing the highly intuitive assessment skills, substantial task experience is indispensable (Klein, 1998). That is, essential to acquiring a sufficient repertoire of situated patterns, a commander has to experience a vast amount of relevant situations. Generally, the common field exercises (FX) fail to provide such practice. FX are generally limited by the features of the dedicated exercise terrains, infrastructure, maneuvering- and weapon-platforms available. In addition, FX are costly, logistically demanding and generally not very efficient as there is always a great deal of waiting…. Worse, crisis situations are generally large scale incidents which are dangerous by nature and are hard if not impossible to mimic properly during a FX.

3 Serious Games for Situation Assessment

Simulation and Serious Games (SG) broaden the range of tasks as well as circumstances under which can be trained and have great potential for the training of situation assessment. Simulations have been here for more than 40 years, but it is the gaming industry that presently defines the progress. Civilian commercial entertainment technology development coincides with the emergence of an impressive military gaming community. The costs of production of content (e.g. terrains, platforms and weapon models) are being shared among the various international military users. As a result, military serious gaming technology undergoes an extremely rapid evolution, more and more commercial off-the-shelf SGs are used in military training courses.

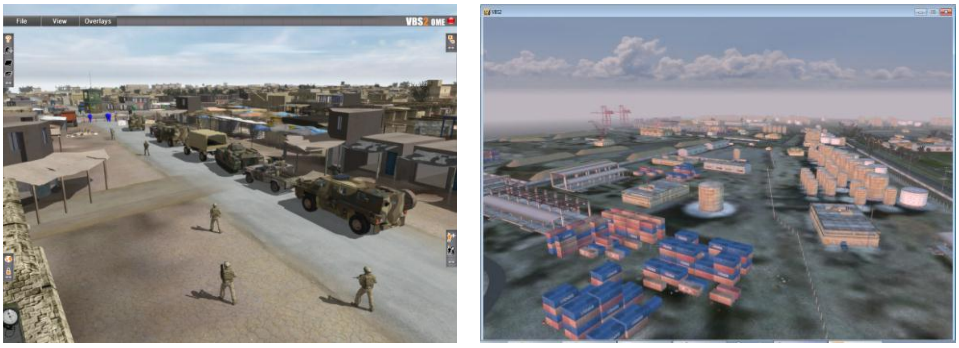

Fig. 1. Tarin Kowt, Afghanistan Port of Rotterdam

Even more important, the recent development of techniques for the (semi) automated generation of terrain databases caters for the fast production of medium to high fidelity 3D geo-specific and geo-typical terrain databases (Smelik et al., 2010). In particularly, the latter are extremely valuable in the development for training in situated assessment.

SGs are currently being used to facilitate e.g. dedicated mission preparation by providing geo-specific representations of mission areas. For instance, many of the primary mission areas of the international coalition forces, Tarin Kowt, Deh Rawod, Kandahar, Helmand and Bagdad have been made available. Such geo-specific terrain databases are used in the international community to train mission-specific situated tactics, to enable troops to gain an understanding of the specific threats at critical locations (infrastructure, high value targets, overwatch locations) and learn to understand the situational tactics, techniques and procedures of the opposing forces (v.d. Hulst et al., 2011b). Also terrains critical to national safety have been modeled, such as the industrial port of Rotterdam, e.g. being used for the training of first responders in handling Chemical Hazards.

SGs will not fully replace FXs, but allow for tasks to be trained that are dangerous, hard, if not impossible, or just too expensive to train at live training sites. SG reliefs the logistic burden of training and allows to reduce time on task and thus allow commanders to experience a great variety of settings within a limited time span. Games are also applied when the use of equipment is too expensive or just impossible, for example in training the tactics of large CAT/convoy operations (v.d. Hulst et al., 2011a). Finally, the nature of SG makes it relatively easy to offer multiple scenarios, either similar or very different, within a short time span. Henceforward, SGs are critical to facilitate the construction of a comprehensive repertoire of patterns needed for RPDM in a variety of circumstances and incidents.

SG do have quite some potential, yet they frequently fail to live up to the expectations as made clear by Hays (2005) who e.g. reviewed 48 empirical research articles on the effectiveness of "instructional" games. Hays is clear in his conclusions; SG Technology alone will not guarantee effective and efficient learning. For SG to be effective, it requires a dedicated training approach.

4 Training Approaches

From the NDM research, several approaches to training complex decision making have emerged. Examples are Decision Skills Training (Pliske et al., 2001) and Job Oriented Training (Stehouwer, 2005, v.d. Hulst, 2008). These approaches provide indications for the instructional design of training for complex decision making and for the implementation of such training.

Decision Skills Training (DST) has been applied extensively in military training for complex decision making, of which, for instance, the application of DST to Urban Operations training for junior leaders (Phillips, et al., 2001) is of particular interest to this work. DST aims at training military to 'learn like experts'. DST strives to provide students with as much relevant experience as possible in the form of a series of increasingly complex scenarios relevant to (aspects) of the decision making. While working in these scenarios, students are trained to mentally simulate possible plans they come upon as a solution to the challenges of the scenario. Also, they are trained to extensively reflect upon their own decision making and to make reflection a habit in their professional life.

To students, DST provides a method for mentally simulating plans, a method for reflecting on the decision making in the scenario's and a method to obtain feedback on the expression of intent.

Job Oriented Training (JOT) is also based on the principles of Naturalistic Decision Making and is a partial implementation and further elaboration of DST.

Personnel in the safety domain is not trained to simply reproduce knowledge, perform standard procedures or solve standard problems. The application of tactical measures must be tailored to best suit the specific mission to be accomplished in a military or safety environment. As a result, training for tactical events should focus on delivering professionals who can act in ever changing and unpredictable situations. In JOT, therefore intention is to target not only conceptual knowledge, but also the skills of independent and competent problem solving in entirely new situations as well as a ‘can do’ attitude, which includes tackling complex situations not previously encountered. The aim, therefore, is to integrate the acquisition of conceptual knowledge, skills and attitude and thus strive for development of rich and integrated competencies.

To target these rich competencies, JOT defines principles for the instructional design of exercises aiming at discovery learning of complex decision making, mostly in a military or first responders context. In JOT, students are confronted with a series of quite short –cyclic, increasingly complex and challenging exercises to allow them discover the essential principles of their job. Crucial is that no theory is provided in advance; theoretical insights are acquired while solving realistic cases. As such, the students do not need to have completed theory oriented training prior to the exercise, as they are expected to discover the essential tactical principles themselves during the JOT exercises. This simultaneously trains them in problem solving in situations entirely new to them and aims at developing a ‘can do’ attitude in tackling new situations. Also, self-reflection is deemed crucial to conceptualize experiences and to make the concepts stick.

Fig. 2. Self – reflection, squad infantry.

Amongst other things, JOT prescribes the nature of instructor support, debriefing and feedback as was deemed crucial by Hays (2005). It also defines requirements with regard to the design of the virtual environments.

JOT was developed for the Dutch Ministry of Defence and has been applied respectively to Serious Gaming for the training of Floodcontrol, Crowd and Riot control (Buiel et al., 2012), Virtual Tactical Trainer for Counter-Improvised Explosive Devices (VTT-C-IED) (v.d. Hulst et al., 2011b), Urban Operations (v.d. Hulst, 2011a), Minewarfare (Stubbe, et al. 2011), Infantry and Cavalry operations (v.d. Hulst,et al., 2008), Naval tactical- and operational tasks (Stehouwer et al., 2006) and Air Defense (Stehouwer et al., 2005). From the above listed efforts in practical applications of the initial JOT concept, many design issues emerged and only few have been solved as of yet.

5 Lessons learned and issues to be solved

The hypothesis underlying JOT was that to obtain optimal effects, the use of SGs for SA demands for 1) a dedicated learning approach as well as a 2) dedicated design of the virtual environment (v.d. Hulst et al., 2008a, 2008b). JOT, therefore, provides high level conceptual prescription for the design of the didactic setting and the design of the virtual environment. Yet, when implementing JOT in realistic settings, many lower level design issues emerge. Those design choices still are found to have substantial impact on the effects of training (v.d. Hulst, et al. 2011a).

Below, we'll list the predominant design issues yet unsolved.

5.1 Fidelity

JOT prescribes that a virtual environment has to provide a 'relevant reality', i.e. a virtual environment that provides the cueing needed to enable adequate SA. The easiest solution to create a relevant reality is to always demand for a high fidelity environment. However, high fidelity models of environments and human behaviour are extremely costly and experience is generally that the costs of such models will exceed available budgets. Also, fidelity studies like those reported in Hays and Singer (1989) provide evidence that, when aiming at tactics, low physical fidelity frequently still yields good learning results. Hence, from a cost perspective, one should aim at defining the minimum level possible that still provides sufficient cueing for SA. In doing so, one must be very careful not to create an environment that leads to negative transfer (Hays & Singer, 1989).

Until now, no generic heuristics to define cueing for SA have been found. Phillips et al. suggest defining cues on the basis of a so called Cognitive Task Analysis, that is, interviewing experts on the cues they use for their SA. Our experience is that it is both hard to find a sufficient number of real experts and time consuming to do a sufficient number of CTAs. Still after defining and implementing the cues, it demands a fair amount of testing to get the cueing right.

Minimum requirements with regard to cueing can well be defined. Visschedijk et al., for instance, used a comparative approach to define the minimum level of cueing needed for proper recognition of emotions in a Crowd and Riot Control training. The authors compared various settings with avatars having either a posture, voice or facial expression representing emotions or combinations of the former. Recognition of those emotions with and without context information was tested with about 20 subjects. This experimental approach, however, is too time consuming to apply at a larger scale. Within a single virtual environment, many different types of cues need to be defined, and such controlled experimentation generally is too expensive during a design trajectory.

Where SA is at the heart of complex decision making, we'll need to find good methods to define the proper level of cueing in Serious Games and adequately test that cueing. Such methods shouldn't be so time consuming that they will not be applied properly.

5.2 Scenario progression

JOT defines that throughout the progression of the training, the complexity of the environment in which the decision making takes place is controlled. The tasks gradually increase in complexity (see e.g. White and Frederiksen, 1990), while performance requirements increase. SA training, therefore, requires a series of scenarios where the initial scenarios are challenging, but do-able and where each subsequent scenario increases in complexity, builds upon the insights acquired in the prior scenarios and introduces sufficient new challenges. In practice, this statement leaves too much to the designer. One needs a good insight which cases are easy and which ones are hard and which ones are really challenging. In designing the military SGs (e.g. the above mentioned VTT C-IED and Air Defence Tactical Training), in the initial phases we had no clue as to the difficulty of cases. We observed that some cases were generally assumed trivial while others were, without exception, found to be really hard. Yet, our Subject Matter Experts had real trouble explaining why cases were easy or hard.

The more SGs we built, we observed some indications that might help to define a good progression of scenarios. We e.g. found out that the easy ones were scenario's that were assumed to be prototypical situations, e.g. situations where all well known factors indicating a threat were present. In contrast, the really challenging cases were those where students had to combine less known indicators with information from several different and sometimes unreliable sources. Also, assumed difficult were the situations very uncommon to the students, e.g. students just back from Afghanistan had trouble imagining the potential threats in a high tech terrorist scenario in a modern city. They were basically looking for the well known threat indicators as known from their Afghanistan experience and didn't look beyond that experience. Still, we observed that there must be many more heuristics underlying the complexity of scenario's, yet still unknown to us.

Without an understanding of factors in complexity of cases, it is impossible to design a good progression. A proper scenario progression is essential to building experience and thus to learning and the field badly needs design heuristics for such progression.

5.3 Flow

The SG world is in a fierce debate about the 'fun' and flow factor (e.g. in Ritterfeld et al., 2009). In our view, Serious Gaming is absolutely not about fun as is, but it is all about flow. That is, the flow that emerges from being fully immersed in a process that is perceived relevant to the job.

If the SG based training is designed well, we see a tremendous flow. In our observations, a major indication for flow is that, while actively engaged on virtual missions, students don't mind extra-curricular hours. When training in a well designed game setting, our students voice no objections to the issuance of orders past 10 p.m. for missions that will start the following day at 7:40 am. and they indicate in their evaluations that this is pretty cool. Their positive attitude towards the training is reflected in their performance, which is generally above standards, as confirmed in the instructor evaluation. Besides effects from using virtual environments for experiential learning, at least time on task is boosted, which is positive in itself since time on task is a one of the predominant predictors for learning (Carroll, 1963).

Fig. 3. Flow- Air Defense Tactical Training.

Design for flow is far from easy. Small inadequacies in the design of the game or the context are observed to make students ‘fight the system’ rather than be immersed.
For instance, we use VBS2 for the creation of several of above mentioned military training environments. When first confronted with VBS2, students automatically try to use it as a first person shooter and start shooting every human being around, just for the fun of it. Only if we design the context (the ‘big game’) right, they refrain from going into ‘entertainment modus’. That is, if we provide a proper briefing, assign roles and provide adequate time for analysis and planning, only then they will use the system to truly train for military operations.
Also, a good design for one target groups isn’t necessarily a good design for another group. We notice in training for the ‘comprehensive approach’ that novices really get into flow and start learning when using the SG Go4it (Hulst, v.d. et al. 2012). In contrast, more experienced students that had been trained to be competitive, however, soon after the initial rounds, just tried to win by playing interventions that were likely to have optimal outcomes irrespective of the validity of those actions at that point in time. Those groups generally do express afterwards that they had fun, but we do observe neither flow nor learning and they quit gaming way before the novices as well as non-competitive experts. We only found out during the trials that we had to provide such a competitive target group with a completely different organisation of the training for it to be effective.
For us, design for flow is to a large extend a trial and error process and indeed we do experience a fair bit of error. Certainly, the field lacks good SG mechanics and design rationales.

Training for Situational Awareness demands for extensive practice and virtual environments allow for such practice and inherently support accelerated learning. SGs hold a great promise for experiential learning especially when the learner experiences the virtual world as emotionally involving and mentally stimulating (Green, 2006).

However, SGs frequently reveal too much error in the design and consequently many games fail to provide such an involving and mentally stimulating environment. Certainly, the design for such environments should become less of an art and more of a science.

6 References

1. Buiel, E., Hulst, van der, A.H. Veldhuis, G. (2012). Inzetmogelijkheden van Serious Gaming voor Public Order Management Training. Soesterberg, TNO Rapport.
2. Carroll, J. (1963). A model of school learning. Teachers College Record, 64, 723-733.
3. Green, M. (2006). Narratives and cancer communication. Journal of communication, 56, 163-183.
4. Hays, R. T. & Singer, M. J. (1989). Simulation fidelity in training system design. New York: Springer Verlag.
5. Hays, R. T. (2005). The effectiveness of instructional games: a literature review and discussion. Technical Report 2005–2004 for the Naval Air Center Training Systems Division: Orlando, FL.
6. Hulst, van der, A.H, Muller, T.J., Besselink, S., Coetsier, D & Roos, C.L. (2008a). Bloody Serious Gaming – Experiences with Job Oriented Training. Proceedings: Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2008.
7. Hulst, van der & A.H., Muller, T.J. (2008b) Job Oriented Training: Onderwijskundige grondslag en onderbouwing. TNO Rapport A194.
8. Hulst, van der, A.H., Muller, T.J., Besselink, S, Smeenk, R. Vink, N. & Visschedijk, G. (2011a). Serious Gaming in Urban Operations: Potential for Tactics Training. TNO Report B073.
9. Hulst, van der, A.H., Amade, C.A.R. & Sain, de, G. (2011b). VTT-C-IED: Threat assessment and planning for search using virtual environments. Nato Research and Technology Organisation - MP-SET-175.
10. Hulst, van der, A.H., Christoph, N., May, A. & Aldershoff, F. (2012, accepted). GCAM: A modeling approach to the Comprehensive Approach. Nato Modelling Simulation and Gaming workshop, Stokholm, October 2012.
11. Klein, G. (1998). Sources of Power, How People Make Decisions. Cambridge, Massachussetts: The MIT Press.
12. Klein, G. (2003). The power of intuition. New York: Random House, Inc.
13. Klein, G. (2008). Naturalistic Decision Making. Human Factors: The Journal of the Human Factors and Ergonomics Society June 2008 50: 456-460.
14. Orasanu, J., & Connolly, T. (1993). The reinvention of decision making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), Decision making in action: Models and methods (pp. 3–20). Norwood, NJ: Ablex.
15. Phillips, J, McCloskey,M. McDermott, P. Wiggins, S., Battaglia D., Thorsden, M.L. & Klein. G. (2001). Decision-centered MOUT training for small unit leaders. Technical Report 1776, U.S. Army Research Institute for Behavioral and Social Sciences, 2001.
16. Pliske, R. M., McCloskey, M. J., & Klein, G. (2001). Decision skills training: Facilitating learning from experience. In E. Salas & G. Klein (Eds.), Linking expertise and naturalistic decision making (pp. 37-53). Mahwah, NJ: Lawrence Erlbaum Associates.
17. Ritterfeld, U., Cody, M. & Vorderer, P. (2009). Serious Games. Mechanisms and effects. New York: Routledge.

18. Schraagen, J.M.C. (1997). Damage Control Decision Support System. In Zsambok, C.E. and Klein, G (1997) Naturalistic Decision Making. Lawrence Erlbaum Associates, Mahwah, NJ.
19. Schraagen, J.M.C. (2008). Naturalistic decision making and macrocognition. Ashgate Publishing, Ltd.,
20. Smelik, R.M., Tutenel, T., Kraker, K.J.D. & Bidarra, R. (2010). Declarative Terrain Modeling for Military Training Games. Int. J. Computer Games Technology. Volume 2010, January 2010.
21. Stehouwer, M., Serné, M. & Niekel, C. (2005). A tactical trainer for air defense platoon commanders. In: Proceedings I/ITSEC 2005.
22. Stehouwer, M., Stricker, J. & Gemeren, van W. (2006). Training Design for Professional Development. In: Proceedings I/ITSEC 2006, Orlando, FL.
23. Stubbé-Alberts, H.E. & Oprins, ET AL.P.B. (2011). Eindrapport JOT-KM: Over de implementatie van competentiegebaseerd leren. TNO-DV 2011 A143.
24. Visschedijk, G.C., Lazonder, A.W., Hulst, van der A.H., Vink, N. & Leemkuil, H. (2012, accepted). Emotions for tactical decision games. Modeling human emotions for tactical decision making games. British Journal of Educational Technology.
25. White, B. & Frederiksen, J. (1990). Causal model progression as a foundation for intelligent learning environments. Artificial Intelligence, 42:99-157.
26. Zsambok, C.E. & Klein, G (1997) Naturalistic Decision Making. Lawrence Erlbaum Associates, Mahwah, NJ.