Selection of a Measurement Battery for Human Behaviour Assessment in Serious Games in the Aviation Domain

G.D.R. Zon

Training, Simulation and Operator Performance department, National Aerospace Laboratory – NLR, Amsterdam, The Netherlands. zon@nlr.nl

The challenge

Is measurement of operator performance behaviour in serious games different from measurement in other environments? A serious game can be seen as an environment in which subjects are allowed and encouraged to play and experiment. The aim is that, by doing so, the subjects learn in an environment that changes, which is safe, and which elicits to play again. As such the serious game can be seen as a mixture of a training environment, a simulation and a game. For each of these domains different ways of measuring behaviour of the human operator exist. The question is to what extent the measurement methodologies that are common for those different domains are also applicable to the measurement of behaviour in a serious game.

Human behaviour assessment in simulator experiments

Normally when assessing operator performance or behaviour in a simulator experiment, the experiment compares a baseline with one or more experimental conditions. From these conditions one or more hypotheses are drawn and compared. These hypotheses may comprise aspects like improved situational awareness, decrease of mental workload, better operator performance, increased trust or operator acceptance, etc. In order to measure those, a number of tools exist. From that toolset a measurement battery for that particular experiment is selected. This is approach is applicable in all kinds of simulation environments ranging from civil or military cockpit to air traffic control environments [3, 7, 8].

Methodological Triangulation

The method "methodological triangulation" is also known as "converging evidence". The method can be summarised by stating that when behaviour of an operator is assessed at least one measure per main category needs to be applied. There are three main categories (see **Error! Reference source not found.**), namely:

- Bio-behavioural or (objective) data (for example psychophysiological or eye tracking data);
- Performance data (for example simulator output like reaction(time) or errors made by the operator);
- Perception or subjective data (for example questionnaires, interviews or workshop outcomes).

The reason why these three main categories are used is that each of these provides information about what has happened during the experiments. However they do that each from their own perspective. A more complete overview of what has truly happened develops when data from the different main categories are compared.

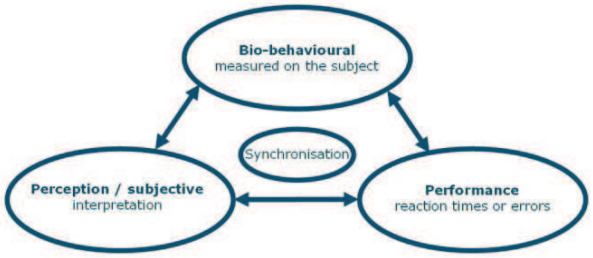

Figure 1. The three main categories that need to be compared and synchronised for methodological triangulation.

Relevant information may be missed when just one or two of the main categories are used. Hence, after participation in an experiment a subject may say that s/he has worked hard and performed well, but are both indeed true? The only way to be certain of that, is by checking other sources of information. Methodological triangulation underlines that: the whole is more than the sum of its individual components.

Serious games

Serious games have an explicit and carefully thought-out educational purpose and are not intended to be played primarily for amusement [1]. They are games that aim to teach the players competencies that are important and relevant for their professional development. Serious games are often used in parallel with other learning tools and environments, such as lectures, e-learning forums, and simulators. They enable a student to 'play' with course material and to see for him/herself how actions that s/he takes may work out without resulting in the consequences that may result from experimenting in real life. They range from very detailed, for part-task training, to very global, to provide students with a good understanding of the coursework. Serious games can even be used as an 'umbrella' for a complete course.

The advantage of using serious games over other learning tools is that when serious games are well-developed, they enhance students' motivation [4] for example because a game is fun to play as well. This increased motivation can lead to students spending more time on the training-task and therefore to better results [6]. In fact the game elicits them to play, and therefore learn, over and over again. A serious game is considered well-developed when the correct balance between entertainment and education was found. Therefore, the development of a serious game is a flexible process with many interactions between developers and potential users to ensure the correct balance. In other words, a game that is both entertaining and ensures real impact (i.e. transfer of learning back into the work-place).

Case studies

Two serious games are currently under development. One within the ADAHR[1] [5] project. This project focuses at the foreseen increase in level of automation in future air traffic control and airport environments. Different roles, or professions, that were identified might be influenced by increased levels of automation in 2035 and 2050. Currently games are being developed in which these new concept can be played with, and tested. There will be two hardware based games in which simulation in realistic environments plays an important role. For a number of relevant aspects that can not be simulated in those environments paper based games will be developed as well. The aim of these games is to acquire better insight in the effect that these higher levels of automation may have on human operators.

The other game is part of the MASCA[2] [2] project. MASCA supports management of airlines and airports in the process of being better able to quickly and swiftly adapt to change. This change can either be induced by authorities but also by market developments or customer requirements. MASCA will create a board, or paper based, game. The game will support the so called collaborative decision making at airports. The players will learn that collaboration between different stakeholders at the airport supports better and faster handing of aircraft. Therefore it will teach them to adjust their behaviour accordingly in the future.

---

[1] ADAHR = Assessment of Degree of Automation on Human Roles (http://www.adahr.eu/ADAHR/). This project is sponsored by the EU and EUROCONTROL in the context of SESAR. The project partners in are: National Aerospace Laboratory of the Netherlands (NLR), Ingenieria de Sistemas para la Defensa de España (Isdefe), Centro de Referencia de Investigación, Desarrollo e Innovación ATM in Spain (CRIDA), Deutschen Zentrums für Luft- und Raumfahrt (DLR).

[2] MASCA = MAnaging System Change in Aviation (http://www.masca-project.eu/). This project is sponsored by the EU. The project partners are: National Aerospace Laboratory of the Netherlands (NLR), Trinity College Dublin (TCD), Kungliga Tekniska Högskolan (KTH), Scandinavian Airlines System (SAS), Swedavia, Aeroporto d'Abruzzo – Pescara, Thales, KITE.

At the moment that this abstract is written the games and as such also the methodology for measurement of the behaviour of the human operators who will be playing the game is under development. Therefore the results and conclusions can not be provided here.

During the presentation

Since a serious game can be seen as kind of hybrid between simulation, game and training, aspects like: operator state (situational awareness, workload, stress, trust, acceptance), operator performance, training competencies that are mastered, feedback, but also relevance for the domain and to what extend the game is fun and elicits to play it more often may need to be assessed. The tools and technologies that were considered, and the set that was eventually selected for this measurement will be presented at Measuring Behavior 2012.

References

1. Abt, C. (1970). *Serious Games*. New York: The Viking Press.
2. Corrigan, S. and Zon, G,D.R. (2011). Managing System Change in Aviation: A New Approach in Human Factors. *Symposium on Human Factors for Future Aviation 2011*. Amsterdam Schiphol.
3. Dijk, H. van, Merwe, K. van de, and Zon, G.D.R. (2011): A Coherent Impression of the Pilots' Situation Awareness: Studying Relevant Human Factors Tools, *The International Journal of Aviation Psychology* **21**:4, 343-356.
4. Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming* **33**(4), 441-467.
5. Koolloos, M.F.J. and Zon, G.D.R. (2011). The Effect of a Highly Automated Environment on Human Behaviour. Plans and first results of the WP-E project "ADAHR". *Proceedings of SESAR Innovation Days 2011*, Toulouse.
6. Mautone, T., Spiker, V.A., Karp, M.R., and Conkey, C. (2010). Using Games to Accelerate Aircrew Cognitive Training. *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*.
7. Zon, G.D.R. and Dijk, H. van, (2009). A selection of Human Factors tools: Measuring HCI aspects of flight deck technologies. *Proceedings of HCI International 2009*. San Diego.
8. Zon, G.D.R., Waard, D. de, and Heffelaar, T. (2008). Human Factors measurement- and analysis tools for cockpit evaluation and pilot behavior. *Proceedings of Measuring Behavior 2008*, Maastricht.