2015

Enroute ATC Industry Perceptions of Simulation Fidelity

Colin Dow

Jonathan Histon

ENROUTE ATC INDUSTRY PERCEPTIONS OF SIMULATION FIDELITY

Colin Dow
University of Waterloo
Waterloo, Canada
Jonathan Histon
University of Waterloo
Waterloo, Canada

Enroute air traffic control (ATC) relies heavily on simulation in training, research, and concept development applications. However, it has little domain-specific research on the effects of simulation fidelity and lacks a standardized definition of simulation fidelity in the literature. A survey of ATC industry professionals shows that simulation fidelity is not perceived to be well defined for the domain of enroute ATC, regardless of respondent nationality, experience, use of simulation or gender. Parts of the operational environment that survey respondents felt were important components in a definition of simulation fidelity are reported; Communications is the most important component regardless of nationality, experience, use of simulation or gender. Implications for the development of a standardized definition of simulation fidelity are discussed.
Simulation fidelity has been researched and investigated for over half a century, yet it remains a somewhat nebulous concept today. The high-level concept of simulation fidelity can be best understood from a definition posited by Hays and Singer: "Simulation fidelity is the degree of similarity between the training situation and the operational situation which is simulated (1989, p. 50)." While this definition is intuitive, more detail is needed for operational applications such as determining the most effective simulation environments for training. Hays and Singer have also provided a more comprehensive definition:

> "Simulation fidelity is the degree of similarity between the training situation and the operational situation which is simulated. It is a two dimensional measurement of this similarity in terms of: (1) the physical characteristics, for example, visual, spatial, kinesthetic, [auditory], etc.; and (2) the functional characteristics, for example the informational, and stimulus and response options of the training situation (1989, p.50)."

Specific components, and more generally a definition of simulation fidelity, are not found in the ATC literature. This is despite the widespread and frequent use of simulation in ATC for a variety of purposes that include training, testing new operational concepts or tools, and research into the future ATC environment. Simulation environments of varying degrees of fidelity are relied upon throughout these different areas, from a classroom-based scenario introducing new trainees to basic concepts to the complex, multi-user MACS simulation used for research at the NASA Airspace Operations Laboratory (e.g. Kraut et al, 2011). When fidelity is reported and discussed within the ATC literature, it is most often in the general terms of low, medium and high; however it is not clear that all the simulators reported in any one category are in fact of equivalent fidelity. For example, Loft et al. (2004) developed an enroute ATC simulation environment intended for research on various human factors related topics. While they discuss their simulator's usefulness as a medium fidelity research tool, the lack of definitions for 'low, medium and high' fidelity make it difficult and potentially ambiguous to compare with other simulation environments . Establishing a definition specific to enroute ATC would provide a formal reference point when discussing simulation fidelity, allowing for critical research to be conducted on the links between simulation fidelity and simulation use within the industry.

Other fields have developed a domain-specific definition of simulation fidelity. For example, Estock et al. (2006) identified and refined specific environmental components (Estock et al. refer to these as dimensions rather than components) that they believe affected the fidelity of a simulation of an F-16 cockpit. Some of the components Estock et al. identified were unique to their work environment, such as the "Visual scene simulation" or "Whole body motion", while others such as "Communications" are important in a variety of work environments. This demonstrates the contextual nature of simulation fidelity definitions, as the components specified by Estock et al. (2006) are appropriate for the simulation of an F-16 cockpit, but many of their components would not work for a simulation of an ATC workstation or an operating room.

An important aspect of the process specified by Estock et al. is that once they had identified their fidelity components, they were verified by consulting with flight simulation experts to determine their validity. As identified by Hays and Singer, receiving feedback from subject matter experts is an important step in defining simulation fidelity for a particular domain (1989). Since they are experts within the domain being studied, their experience with the operational environment will be able ensure that no components have been overlooked.

This process of narrowing the focus of a fidelity definition to be highly domain specific is necessary for researchers to be able to study how fidelity is perceived in a given work environment. More importantly, this allows for objective research into the links between fidelity and simulation use for training, testing new operational concepts and research within the given domain. Developing a clearer picture of what components affect fidelity for a particular operational environment opens up the potential for using a variety of different simulation environments to achieve outcomes in each of these areas in perhaps a more effective and cost-efficient manner.

As part of a project developing a simulation fidelity definition for enroute ATC (e.g. similar to the Estock method note above), an industry wide survey was conducted investigating the perceptions of simulation fidelity and how simulation of varying degrees of fidelity ought to be used. Considering that the process of defining simulation for a particular domain has seldom been done, and never for enroute ATC, there was a clear opportunity to develop a domain-specific definition of simulation fidelity for enroute ATC (Dow and Histon, 2014). Included in this survey were questions that sought to determine if simulation was already a well defined concept within ATC, and what environmental components individuals were currently considering when making a fidelity determination for a simulation environment.

As part of assessing the potential for a general consensus on the appropriate components in an ATC simulation fidelity definition, this paper compares the perceptions of different demographic sub-groups regarding the need for a standardized simulation fidelity definition for the ATC domain, as well as the particular environmental components that they believe ought to be considered when defining simulation fidelity for enroute ATC. The demographic sub-groups used for comparison are nationality, survey participant's primary use of simulation, survey participant's years of experience with simulation, and gender. What follows is a description of the methods used to gather the professionals' perceptions on simulation fidelity, presentation and analysis of the results, and finally a discussion of the findings and limitations of the study.

Method

As part of a larger effort investigating the concept of simulation fidelity within the ATC domain, professionals within the ATC industry were asked a series of questions about simulation fidelity through an online survey. This paper focuses on the responses to two questions in the survey. The first was a question on whether or not professionals believe simulation fidelity is well defined within the enroute ATC industry. The second question, asked individuals to provide a list of the environmental components that they believed affect the fidelity of a simulation environment.

The survey was first distributed to personal contacts within various ANSPs and researchers around the world who met the participant criteria of the survey. This was done to try to ensure that the survey participants were coming from as reliable a source as possible due to concerns about the lack of control and verifiability of those completing an online survey anonymously. The target population was anyone who had experience developing or using ATC simulations, which included the following examples of potential participants:

- Active air traffic controllers who have used simulation for training / participated in simulation studies
- Controller training designers / developers
- Air traffic control instructors
- Researchers who have used simulators for human-in-the-loop studies
- Operational concept developers and controller tool developers who have used the results of simulation studies

The survey was then made publicly available on aviation public domain websites (e.g. liveatc.net, pprune.org) and through air traffic control publications (e.g. ATC Network and Air Traffic Management) where the target population for this research typically frequent. Free response questions provided opportunities to carefully screen responses for appropriateness and consistency with the background and experience reported by the participants.

Survey questions consisted of a mix of Yes/No, Likert scale ratings, and short and long answer questions. Topics covered in the survey include participant perceptions regarding the concept of simulation fidelity in the domain of air traffic control, what level of simulation fidelity is required to train for a certain skill or test/evaluate a particular concept, and the acceptability and accuracy of a simulation fidelity definition and categorization system developed for the enroute ATC domain. Full results are presented and discussed in Dow (2015).

There were 86 completed responses. The key characteristics of participants were gender (Male=69, Female=16, Prefer not to respond=1), nationality (United States=34, Canada=29, International=22, Not specified=1), years of experience working with simulation (0-5 years=23, 6-10 years=11, 11-15 years=14, 16 years or more=38), and the survey participant's primary use of simulation (Training=58, Testing new operational concepts=17, Future ATC environment research=11).

Results

Participants were explicitly asked whether or not they believe simulation fidelity was a well-defined concept in the domain of ATC. Participants were given radio buttons and could chose either Yes or No and were asked to explain their choice in a free-response text box. The results are presented in Figure 1.

The 'Yes' and 'No' columns in Figure 1 represent responses where participants demonstrated a clear understanding of the question based on their follow-on explanation of why they answered Yes or No. Survey participant explanations that clearly indicated they did not understand the question were categorized as 'Non pertinent'. For example, a response of, "I have been working in ATC for 23 years, and simulation has been in use all of this time." was judged to indicate the participant had not understood the question. The final column, 'No explanation', represents the percentage of responses where survey participants provided no explanation to their Yes/No answer and therefore an assessment of their understanding of the question could not be established. The analysis below focuses only on the Yes and No response columns, referred to henceforth as the "analyzed responses" in the subsequent analysis.

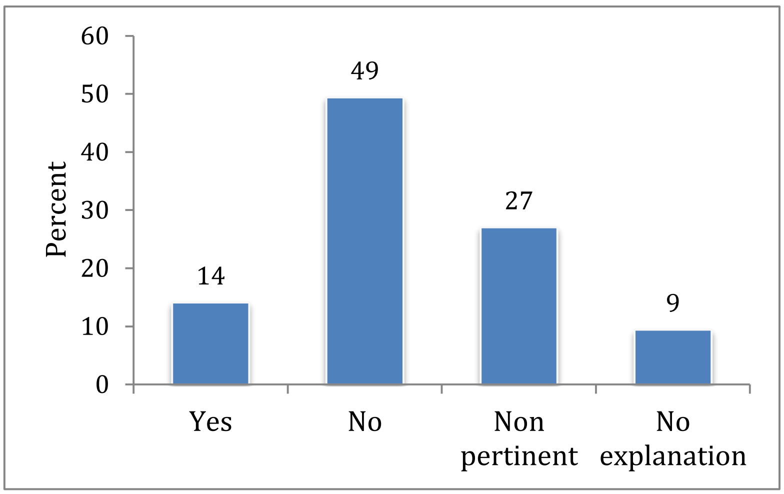

Figure 1. Survey participant responses to the question: "Do you believe that simulation fidelity is a well-defined concept in the ATC domain?" N=85.

Results of a chi-square goodness of fit test show that the observed Yes/No response frequencies are statistically significant. They are different from what would be expected if half of the ATC industry believed that simulation fidelity is well defined for the ATC domain (χ^2(1, N=54)=16.67, p<0.001). Demographic data collected as part of the survey was used to investigate whether the perception that fidelity is not well defined is wide-spread across gender, nationality, experience and primary use of simulation. A comparison between the Yes/No response rates for these four demographics is presented in Figure 2. As seen in the figure, the proportion of Yes/No responses, while varied, shows a strong and consistent pattern of a belief that simulation fidelity is not well defined. A chi square analysis was performed to determine if there were any differences within the demographic groups. It was found that there were no differences with regards to the belief that simulation is **not** well defined for ATC when comparing within the demographic groups of gender, (χ^2 (1, N=54)=0.04, p=.851), nationality, (χ^2 (2, N=54)=2.06, p=.385), years working with simulation, (χ^2 (3, N=54)=3.78, p=.287), or survey participant's use of simulation, (χ^2 (2, N=54)=1.83, p=.400.

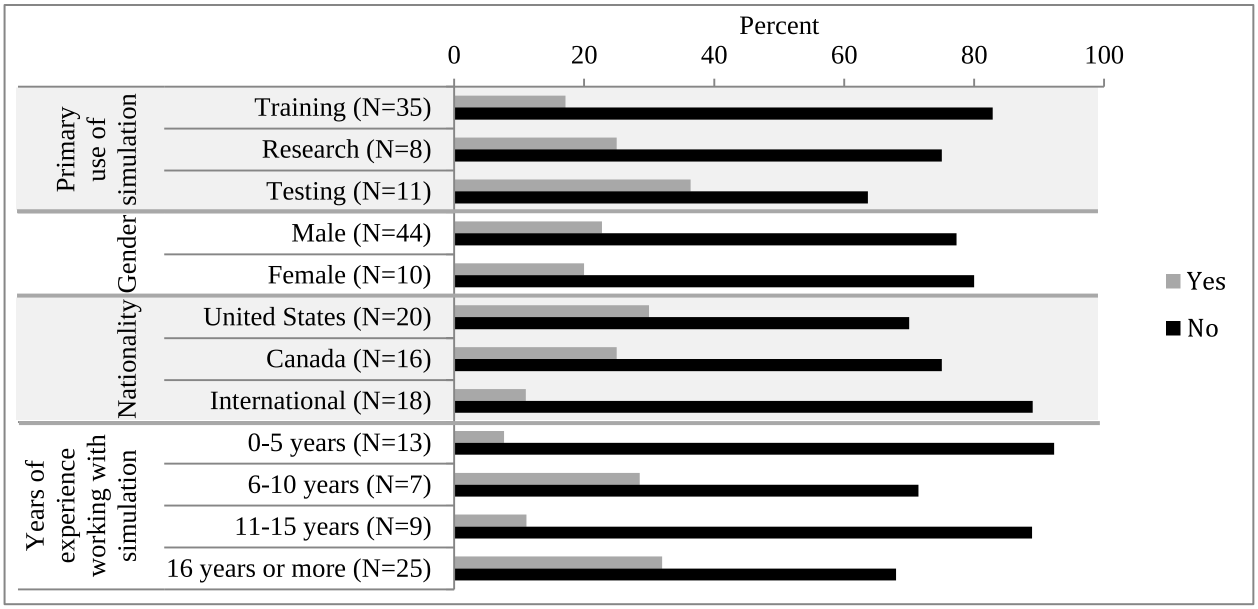

Figure 2. Demographic sub-group comparison of responses to the question: “Do you believe that simulation fidelity is a well-defined concept in the ATC domain?” N=54

In order to further explore why survey participants feel simulation fidelity is not well defined for ATC, Table 1 presents sample comments from both the pertinent “Yes” and “No” responses to the follow-up question asking if they could explain their answer in more detail.

Table 1.
Sample comments from survey participants’ explanations of their responses to the question in Figure 1.

Sample Comments from ‘Yes’ Responses	**Sample Comments from ‘No’ Responses**
• I think that it is well defined, but in reality, it is under-utilized. • We all know what fidelity means. Realistic. Realistic aircraft, realistic routes, realistic responses. Responses that are dynamic in nature, changing depending on what the student is doing. • Though I'm not aware of a quantitative definition, fidelity is something researchers and trainers know when we see it, and it is easy to ordinally rank different simulators or simulations in terms of their fidelity. I have created and used an informal table that lists the different levels of fidelity and their characteristics. • I think it's defined and conceived just fine, but, in my opinion, it's not implemented very well.	• I think that "simulation fidelity" is one of those concepts that "everyone knows what it means" but that formal, valid definitions are lacking. • I believe simulation fidelity means different things to different people. I believe current controllers are not involved enough in validating the fidelity of a simulation before it is used in the field. • I have not come across such a concept definition so far. On the contrary, many times the term "high fidelity" is interpreted in various ways. • I’ve met a lot of people in my business who have a significantly different perception of what is high and what is low fidelity simulation. • My interpretation of high fidelity simulation is the recreation of the real live ATC environment in as much detail as possible. I don't believe this to be a universally shared interpretation and that there are varying degrees of separation from my idea.

The sample comments from those who responded “Yes”, are representative of a recurring belief that simulation fidelity is a well-defined concept, but is not put into practice or referenced enough with regards to the many uses of simulation within the industry. However, what is clearly demonstrated in the sample comments from those who responded “No” is that the problem is not with an individual’s definition in isolation, but rather when discussing the issue as a collective and not sharing the same definition with those they interact with.Comments such as “I believe simulation fidelity means different things to different people”, “On the contrary, many times the term "high fidelity" is interpreted in various ways”, or “I don't believe this [his/her interpretation of fidelity] to be a universally shared interpretation and that there are varying degrees of separation from my idea", all indicate an awareness of the impact of a lack of standardization with regards to simulation fidelity in the ATC domain.

In addition to the question above, survey participants were asked to provide the environmental components they believed affected the fidelity of a simulation environment. Eight optional text boxes were provided to participants in order to receive as many different environmental components as possible. The responses were then coded by identifying a high level topic or theme in the response, and the top ten coded reponse frequencies from the overall responses as well as the nationality demographic groups are presented in Table 2.

Table 2.
Top ten coded response frequencies for all survey participants and nationality demographic groups for the question "In your opinion, what parts of the enroute ATC work environment affect the fidelity experienced by someone using an enroute ATC simulation?"

Fidelity Components	**Response frequency (% of N)**			
	Overall (N=73)	**United States (N=31)**	**Canada (N=24)**	**International (N=18)**
Communications	62	55	71	56
Equipment	42	35	46	44
Environment	32	32	42	17
Aircraft performance	30	16	46	28
System participants	29	23	38	22
Unpredictability	29	19	42	28
Traffic	23	19	25	28
Weather	21	10	42	11
Automation	19	16	17	28
Operational stress	11	6	13	11

Tables similar to Table 2 were also prepared for the demographic groups of gender, simulation use, and years of experience with simulation; however, they are not shown due to space considerations. Across all demographic groups, the "Communications" component received the highest response frequency for each sub-group, indicating its high overall rank was the result of a widespread and shared perception of its importance for a definition of fidelity for enroute ATC simulation environments. For a detailed description of the fidelity components, see Dow (2015). Not all components appear to be perceived equally across the different nationality groups, though statsitical tests of significance have not been completed. For instance, Canada had a much higher response frequency for "Weather," while the United States had lower response frequencies for "Aircraft performance", and the International group had lower response frequencies for the "Environment" component but higher response frequencies for "Automation". From the tables not shown, the researchers demographic group overwhelmingly identified Communications (73%) and Equipment (45%) components, while all other components were at less than 27%. The demographic group of testing new procedures had almost no (< 7%) mentions of Unpredictability, Weather, Automation, and Operational Stress. Table 2 also illustrates that there were differences in how many components each nationality was providing, with Canadian survey participants providing more components then the other two groups.

Discussion

The results presented in Figure 1 showed that simulation fidelity is viewed as not being well defined for enroute ATC; this indicates that there is an opportunity for developing a standardized definition of simulation fidelity for the enroute ATC domain. The consistency of this finding within the different demographic sub-groups in Figure 2 suggests that the notion of simulation not being well defined is wide spread across gender, nationality, survey participant's simulation use, and their years of experience with simulation.

The examination of what components participants felt contributed to simulation fidelity indicates some potential sources of this perception, as well as the basis for development of a standardized definition. It is clear that the components listed in Table 2 are not unanimously agreed as only one component (Communications) was identified by more than half of all participants. The response frequencies from the different nationality groups also showed that it appears individuals are considering different parts of the operatinal environment when making a determination regarding the fidelity of an enroute ATC simulation environment. While the overall response rates give confidence in drawing conclusions for the primary results, the small number of participants (minimum of 11) within some of the demographic groups suggests caution in interpreting the findings for components for any one sub group. The presence of the differences, however, is consistent with the ambiguity and confusion around the concept of simulation fidelity and the difficulty in having objective discussions regarding the implications of simulation fidelity, let alone conduct research regarding the link between fidelity and training, for example.

However, the components listed in the first column of Table 2 also offer a reasonable consensus of the components that can affect the fidelity of an enroute ATC simulation environment. Identifying the range of components that are considered and capturing them within a proposed definition of simulation fidelity (see Dow and Histon, 2014) is an important step towards developing an operationally useful and widely accepted understanding of simulation fidelity in enroute ATC.

Summary

Given the significant amount of survey participants who believe simulation fidelity is not well defined for ATC, and the variation inherent with individuals' sets of fidelity components, there is an opportunity for increased standardization by developing a definition of simulation fidelity for the enroute ATC domain. One such definition has been developed by Dow and Histon (2014). This construct presents a set of environmental components that can affect the fidelity of an enroute ATC simulation simulation environment. Most important is the process it has also taken to include SMEs in both the development of the construct and the validation of the final product. As noted earlier, this process of validation by those who work closest with the operational environment being simulated is important in not only developing a definition that captures the relevant components, but one they will have more confidence in once they are using it.

Even if a definition were developed, this does not close the door on the topic of simulation fidelity. It is a first step to a clearer understanding of the concept, and more work is needed. What a definition will provide is the foundation of how simulation environments are compared and contrasted, essentially the points of comparison. The definition would then need to be operationalized in some form to be able to clearly delineate between simulation environments. The most likely form is that of a categorization system similar to that used by the FAA to classify flight simulation environments but for enroute ATC simulations. This preliminary work then creates the opportunity for the important research in to the link between fidelity and simulation use for training, testing new operational concepts and future ATC environment research.

Acknowledgements

We would like to thank the National Sciences and Engineering Research Council (NSERC) for their financial support of this project.

References

Dow, C. and Histon, J. (2014). An Air Traffic Control Simulation Fidelity Definition and Categorization System. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting,* 58(1), 92-96. DOI: 10.1177/1541931214581020

Dow, C. (May, 2015). *Developing an Objective Definition of Simulation Fidelity for Enroute Air Traffic Control.* (Master's dissertation). University of Waterloo; Waterloo, ON, Canada.

Estock, J. L., Alexander, A., Gildea, K. M., Nash, M., and Blueggel, B. (2006). A Model-based Approach to Simulator Fidelity and Training Effectiveness. In *Proceedings of Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC).* Orlando, FL: National Security Industrial Association. Retrieved from design-usability.aptima.com/publications/2006_Estock_Alexander_Gildea_Nash_Blueggel.pdf

Hays, R.T. and Singer, M.J. (1989). *Simulation Fidelity in Training System Design: Bridging the Gap Between Reality and Training.* New York, NY: Springer.

Kraut, J., Kiken, A., Billinghurst, S., Morgan, C., Strybel, T., Chiappe, D., Vu, K. (2011). Effects of Data Communications Failure on Air Traffic Controller Sector Management Effectiveness, Situation Awareness, and Workload. *Human Interface and the Management of Information. Interacting with Information Lecture Notes in Computer Science*. Volume 6772, pp 493-499. DOI: 10.1007/978-3-642-21669-5_58

Loft, S., Hill, A., Neal, A., Humphreys, M., & Yeo, G. (2004). ATC-lab: an air traffic control simulator for the laboratory. *Behavior Research Methods, Instruments, & Computers*, 36(2), 331–8. Retrieved from http://www.ncbi.nlm.nih.gov/plubmed/15354699