Training, Education, and Learning Sciences, Vol. 193, 2025, 20–30
https://doi.org/10.54941/ahfe1006642

Developing a Post-Flight Debriefing Tool: Insights From a UX Workshop and Tool Demonstration

Maria Hagl[1], Darlann Bertolone-Lopez-Serrano[1], Marcus Biella[1], Alexander Donkels[2], Andreas Volkert[1], Sven Lorenz[2], Kristian Fettig[2], Nicholas Sloane[2], and Andreas Pick[1]

[1]German Aerospace Center DLR, Institute of Flight Guidance, 38108 Braunschweig, Germany
[2]German Aerospace Center DLR, Institute of Flight Systems, 38108 Braunschweig, Germany

ABSTRACT

After flight training, student pilots and flight instructors debrief by discussing the trainee's performance and sharing observations. To improve this process and promote more objective debriefing, a web-based debriefing tool was developed to integrate and present in-flight data (e.g., vector data and position, etc.) and operator data (e.g., physiological and psychometric data, etc.). The tool allows users to replay critical flight segments for improved flight training analysis for, but not exclusively, maritime drone crews. Its development followed a user-centered design approach, with insights gathered through a user experience (UX) prototyping workshop and follow-up tool demonstration with experimental flight crews (both in the role of trainees and trainers) in unmanned aviation. Participants evaluated the tool's user experience before and after the tool customization. While the debriefing tool is perceived as rather neutral, participants embraced the concept and gave valuable insights on effective debriefing and on improving the tool. In general, we can state that the flexible design allows easy customization for diverse user needs, offering a promising framework for effective and more objective debriefing across various domains in aviation.

Keywords: Training, Debriefing, Learning, Debriefing tool, UX, User-centered design

INTRODUCTION

Background on Tool-Supported Debriefing

Debriefing is an important part of flight training that occurs at the end of a training session to discuss the trainee's performance and share observations. Whether conducted individually or as a team, debriefing guides participants through a set of questions that enable them to *"reflect on a recent experience, construct their own meaning from their actions, and uncover lessons learned in a nonpunitive environment"* (Tannenbaum and Cerasoli, 2013, p. 231). In aviation, instructors and their students typically debrief within an hour after training, when impressions of the exercise are still recent (Wiese et al., 2008, p. 290). Intelligent mission debriefing tools were already developed for military aviation during the 1980s (see Snipes, 1989). Since the early to mid-2000s, debriefing in aviation after simulator training has increasingly benefited from technological assistance (see Allen et al., 2004; Burns et al., 2003; Jensen et al., 2006; Schmorrow et al., 2008, p. 213). More precisely, well-engineered computer-supported debriefing tools support instructors to gain a clearer overview of participant performance, even under existing time constraints (Wiese et al., 2008, p. 289). According to Greschke and Cerutti (2003, p. 7), "*for debrief, the ability to fully reconstruct the mission, enabling replay of any mission event, is desired*". During training, instructors can mark significant events in certain training simulation platforms and replay the session, helping trainees to revisit their strengths and errors (Wiese et al., 2008, pp. 290–291). Serial replay (Wiese et al., 2008, p. 290) and revisiting training elements could enhance learning by allowing student pilots to directly relive elements from the training, rather than relying on memory of the positive or negative actions observed by the instructor (Roth, 2015). Current tools supporting debriefing focus on specific methods such as eye-tracking (Ryffel et al., 2019), while others search to combine a more comprehensive set of relevant training data (Bunn et al., 2020) or even rely on artificial intelligence (Henquet and Bellucci, 2024).

Developing a Debriefing Tool for Maritime Drone Crews

At the Institute of Flight Guidance of the German Aerospace Center, the Human Factors department (German Aerospace Center, 2025) is developing a debriefing tool to make after-action reviews following human factor studies more objective and effective. As part of its scientific activities, the department regularly conducts experimental research involving various aviation stakeholders, such as pilots and air traffic controllers. To successfully apply the debriefing tool after human factor studies, its development must focus on flexibility and modularity, allowing it to be easily customized to various use cases and user groups.

This article addresses the customization of the debriefing tool for crews operating remotely piloted aircraft systems (RPAS) in a maritime environment. As part of the Maritime RPAS 3 project (MaRPAS 3, German Aerospace Center, 2024), a new training concept has been developed for future drone crews in maritime operations, explicitly addressing the needs of Germany's Federal Police. More specifically, the development of the debriefing tool was centered on creating a web-based application to support more objective and effective training debriefings of a maritime drone crew. This application will assist operators, such as drone pilots and instructors, in visually reviewing training elements in debriefing by displaying relevant data, including flight metrics and operator-related statistics.

To ensure the application is user-friendly for drone flight training missions, we followed a user-centered design approach (International Organization for Standardization, 2019). The requirements for the debriefing tool were defined during a prototyping workshop that followed a design thinking approach (see Lewrick et al., 2020, p. 19), and integrated into the development process over five months. After its first implementation in a human factors study, the tool was re-evaluated to assess the extent to which it met practical demands.

General Information on the Developed Debriefing Tool

The debriefing tool is a web-based application, programmed using JavaScript and Python. The client-side module provides an intuitive interface that allows users to interact with the tool easily. Users can upload training-related files after selecting a predefined data type (e.g., flight data, electrocardiography data, etc.). The data are transmitted to the backend, where all data processing and storage aspects are handled. Once processed, the data can be visualized in various formats, including maps and graphs.

The tool is organized into relevant tabs, such as ‘Flight Data’, ‘Operator Data’, and more. Additionally, each page is fully customizable, allowing users to configure and arrange windows and display only the specific data they wish to visualize. **Figure 1** illustrates the debriefing tool as a mock-up in its prototype version, while **Figure 2** shows the tool in its current development stage.

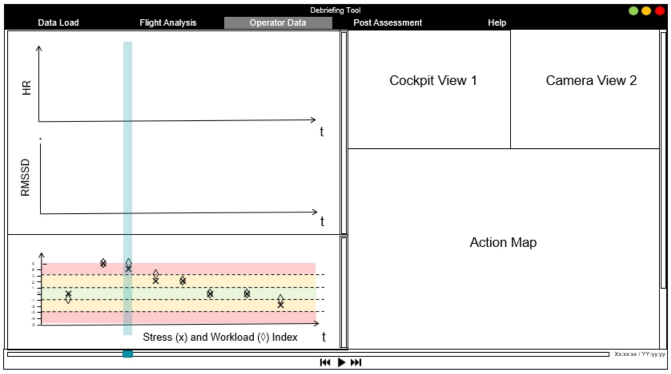

Figure 1: Mock-up of the debriefing tool in its prototype version (May, 2024).

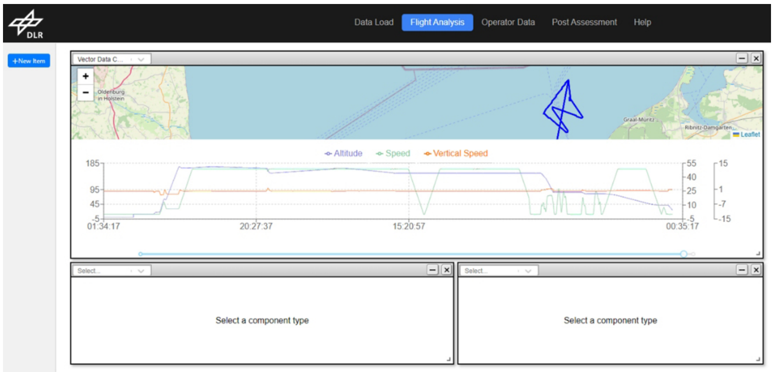

Figure 2: Debriefing tool in its current development stage (October, 2024).

METHOD

To customize the debriefing tool to the requirements for training maritime drone crews, we conducted a three-hour prototyping workshop as the first part of the study. The second part of the study involved a demonstration of the debriefing tool, followed by its evaluation as part of a human factors study conducted to assess a training concept for maritime drone crews.

Sample

In both parts of the study, the sampling method was purposive sampling. The participants in the prototyping workshop were five experts with varying degrees of experience across multiple roles in experimental drone flight operations. These roles included operating a ground control station, manually piloting drones, managing sensors attached to drones, and serving as trainers and trainees. Six air vehicle operators and six sensor operators ($N = 12$) evaluated the tool following a demonstration during a human factors study's debriefing session. Only eight participants submitted fully completed questionnaires, while four omitted the questions of the reverse side. Consequently, we assume the data to be missing completely at random and exclude the partially completed questionnaires from the analysis. While this further reduces the already low sample size, the impact is estimated to be minimal given the study's small scale.

Procedure and Material

The prototyping workshop was led by two facilitators and began with an open brainstorming and mind-mapping exercise on flight training to introduce participants to the topic. The next step involved an adapted version of the sailboat exercise (Bordogna, 2018), which aimed to identify the positive and negative aspects participants associate with debriefing after a training session. Participants first worked individually with post-its: blue was used to note positive aspects, and orange for negative aspects. These were then shared by each participant and presented by the facilitators. The facilitators categorized these aspects on a whiteboard using keywords to identify the main themes. Next, the first prototype version of the debriefing tool (see **Figure 1**) was introduced. After the participants had time to absorb the design and ask questions, they individually completed the user experience questionnaire (UEQ, see Schrepp, 2015; Schrepp et al., 2017; Schrepp and Thomaschewski, 2019). The workshop then continued with a rose-bud-thorn exercise (see Stackowiak and Kelly, 2020, pp. 61–62), where participants used yellow post-its to identify positive aspects, pink for challenges, and green for points of improvement of the prototype. These points were then shared by participants and grouped by the facilitators on boards. After a brief break, participants were given printed versions of the mock-up to sketch their ideas, which were then presented to the group.

The debriefing tool developer incorporated the generated ideas as much as possible, further refining the tool (see **Figure 2**). Five months later, the updated version of the debriefing tool was presented to participants in the second part of the study. Six experimental drone crews ($N = 12$) could try the tool and evaluate the UX via the UEQ.

RESULTS

This section first presents key insights for effective debriefing, followed by initial reactions to the mock-up of the debriefing tool prototype. We will then show results regarding the user experience during the workshop and after the demonstration of the customized tool.

Key Insights for Effective Debriefing

The keywords from the adapted Sailboat Exercise allowed us to generate five main topics for effective debriefing (a-e), which are outlined below in detail.

(a) *Keep the delay between training and debriefing short*

Minimizing the time between a training flight and the debriefing is perceived to be essential for participant engagement and retention while the experience is fresh. As a rough timeline, one participant suggested that the delay between training and debriefing should not exceed 30 minutes. Additionally, another participant highlighted the importance of allocating sufficient time for debriefing to address all relevant training aspects thoroughly.

(b) *Debriefing should occur within a structured framework*

For workshop participants, debriefing should follow a structured framework to ensure all involved parties can prepare and follow through. One participant suggested a neutral facilitator take notes throughout the training session and lead the debriefing afterward. An observation log is considered crucial to ensure that all relevant aspects of the training are captured, as a lack of more objective data can negatively impact the debriefing process. For instance, discussions focused solely on personal beliefs and experiences without objective support are perceived to be counterproductive.

(c) *Promote a democratic and equitable approach to debriefing*

Participants in the workshop emphasized the need to uphold equality and adopt a democratic approach during the debriefing. They believe it is important to maintain a flat hierarchy, allowing everyone to share their views without being interrupted. The proposed moderator is essential to engage all parties equally to ensure a variety of perspectives are presented. It is considered important to create a space where subjective impressions, ranging from observations to feelings, can be shared openly without drifting into side discussions.

(d) *Create a supportive space for comfort, trust, and focus*

Ensuring a supportive space for debriefing is essential for workshop participants. They emphasized that having a comfortable and private setting helps to focus and facilitates communication without distractions. The atmosphere should promote trust by establishing a non-judgmental space that encourages self-reflective questions and open discussion.

(e) *Agree to disagree and end the debriefing on a positive note*

According to workshop participants, debriefing should follow the motto *"agree to disagree"*, allowing for conflicting perspectives while avoiding debates on right and wrong. This requires discipline from all involved parties to avoid placing personal blame during discussions. The debriefing should provide a comprehensive review of performance and identify areas of improvement. In addition to lessons learned, successful training elements should be highlighted to ensure trainees leave on a positive note.

Initial Reactions to the Mock-Up of the Debriefing Tool Prototype

The initial reactions to the mock-up of the debriefing tool prototype developed during the Rose Thorn Bud exercise, can be clustered into five general topics (f-j) described in detail below.

(f) *Facilitate objective debriefing without making the tool its central point*

Overall, it was positively noted that the tool intends to support the team with objective data during the debriefing process. This would allow the team to rely on concrete information rather than on thoughts and opinions. However, there was criticism regarding the potential loss of valuable face-to-face discussion post-flight if the tool became too central to the debriefing process.

(g) *The tool's conception might serve more as an expert tool post-debriefing*

Participants in the workshop remarked on the need to synchronize a substantial amount of data from various sources before the debriefing session. This requirement could extend the time between the flight training and the debriefing session. Additionally, several participants questioned the tool's relevance for operators during the debriefing, suggesting it might be better suited as a complex flight analysis tool for expert evaluation after the debriefing.

(h) *Envision the tool as a cloud-based collaborative platform*

Workshop participants discussed the potential of using the tool as a cloud-based collaborative platform, enabling multiple team members to access it simultaneously. Another positive feature highlighted was the tool's ability to collect data for future reference. Additionally, it could be designed as a learning tool where user profiles are created for both trainers and trainees, allowing them to monitor their learning progress.

(i) *The overall design is well structured but lacks pictograms and colors*

Workshop participants evaluated the prototype's design as clearly structured and well laid out, with the flexible tab and window structure being rated particularly positively. However, one potential drawback mentioned was the overwhelming amount of data that could be displayed at once. The small font sizes and the lack of pictograms or colors were perceived as rather negative.

(j) *The prototype needs improvement, particularly in its special features*

Several features were suggested by the workshop participants to improve the tool's functionality. One user recommended adapting the tool for mobile applications, which is perceived particularly beneficial for flight training outside the simulator. Other suggestions included the ability to add comments or bookmarks in flight data, logs, and cockpit presentations. Additional requests involved more options for time frame management (e.g., frame-by-frame or interval view) and replay speed control. An export function for use in papers and internal reports, along with the option to load different post-assessment onsets, were also mentioned. Lastly, there were requests for drop and drag menus and resizable windows.

User Experience in the Workshop and After the Tool Demonstration

The UEQ results from the workshop show a neutral evaluation of the debriefing tool across its dimensions of attractiveness, perspicuity, efficiency, stimulation, and novelty, while dependability received a positive evaluation. The data indicate low agreement among participants for perspicuity ($M = .00$; $SD = 1.36$), attractiveness ($M = .17$; $SD = 1.2$) and efficiency ($M = .25$; $SD = 1.13$), medium agreement for novelty ($M = -.35$; $SD = .95$), as well as high agreement for dependability ($M = .95$; $SD = .74$) and stimulation ($M = .15$; $SD = .68$).

After the demonstration of the debriefing tool, it is evaluated neutrally across all dimensions of the UEQ. The data show medium agreement among participants for perspicuity ($M = .41$; $SD = .97$) and dependability ($M = .53$; $SD = .91$), as well as high agreement for attractiveness ($M = .23$; $SD = .68$), stimulation ($M = .41$; $SD = .67$), efficiency ($M = .63$; $SD = .55$), and novelty ($M = -.28$; $SD = .53$).

When comparing the user experience results from the prototyping workshop to those from the follow-up tool demonstration (see **Figure 3**), the user experience is overall neutral in both cases. While the differences are not significant for any of the dimensions, there is a slight trend toward improvements in perspicuity, efficiency, and stimulation when moving from the prototype to the customized tool. However, there is no noticeable change observed in terms of attractiveness and novelty. On the other hand, results show that dependability has decreased from the prototype to the customized tool.

DISCUSSION AND CONCLUSION

The current version of the debriefing tool was designed with sufficient flexibility to be customized according to the needs of different user groups and use cases – in this case, to meet the potential requirements of a maritime drone crew as part of the MaRPAS 3 project.

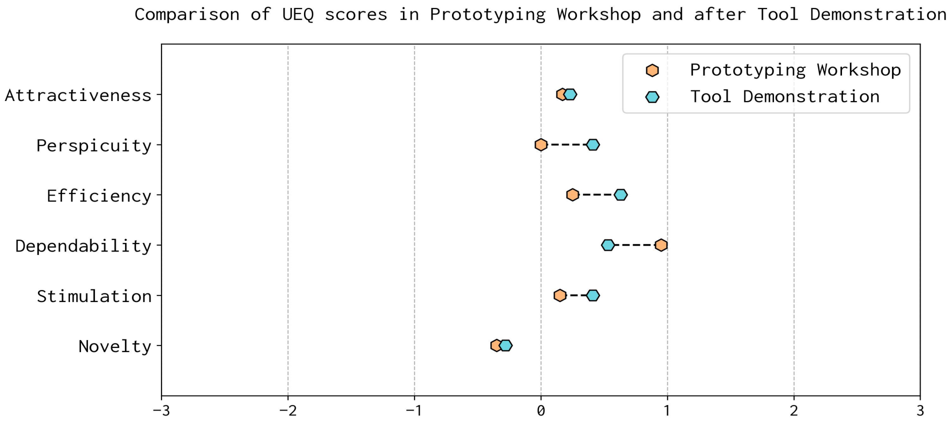

Figure 3: Comparison of UEQ scores in the prototyping workshop and after the tool demonstration.

From the prototype to the customized tool, the hedonic quality of the product remains largely unchanged. While the debriefing tool appears to be slightly but insignificantly more engaging and motivating to use (stimulation), there is no noticeable difference in the creativity of the design. This lack of novelty is not surprising, as the general layout of the customized tool is strongly inspired by the mock-up prototype version. The overall perception of the tool, both in its prototype stage and as a first developed version, is perceived as nearly equally attractive and largely neutral. This result is acceptable, as workshop participants highlighted that the tool should be able to facilitate a more objective debriefing without becoming the focal point of attention. In other words, the tool should play a rather functional role during debriefing. Regarding the pragmatic quality of the current system, there are insignificant but noticeable differences between the prototype and the customized tool. The developed version seems to be easier to understand (perspicuity), allowing users to operate it with less effort (efficiency). Furthermore, there is a slight decline in the feeling of control over the interaction with the tool, making it seem less secure and predictable (dependability). This result can be explained as follows: while the tool is intuitive to use, there are still synchronization problems that question the dependability of the system. In the human factors study, the tool was not well integrated into the simulation environment. The predefined data sets came from various sources and formats and there was no effective time synchronization among them. While the lack of a common start time is not a problem of the debriefing tool per se, the replay function was only effective for individual datasets and could not be synchronized across multiple datasets as initially intended. This problem had been identified before the human factors study and despite it being a desired feature by both the developers and the workshop participants, no general slider was implemented within the short development time. Additional issues arose in the display of times across the data visualization: While some data sources record a starting point from the beginning of the simulation, others log the local time. For a successful implementation of the debriefing tool in the future, it is crucial to ensure that simulation data share the same timestamp and converge at a centralized node within the network to prevent synchronization discrepancies. Furthermore, the debriefing tool should identify the common starting point of the training run as the zero reference and automatically translate it into simulation time to eliminate display inconsistencies.

A limitation in comparing the perceived user experience between the prototype and the customized tool is that workshop participants had to imagine how they would interact with the tool, as they could not set hands on the mock-up prototype. Nevertheless, a positive outcome was that, on average, users did not perceive the tool negatively during the workshop or in the follow-up evaluation. Furthermore, no participant in either phase of the study outright rejected the tool.

Despite the overall acceptable results for the applied use case, there are several areas for improvement that must be considered for the further development of the debriefing tool. First, the categorization of the data must be automated to ensure efficient data transfer and, consequently, minimize the time between training and debriefing. Second, predefined and storable visualization setups are necessary so that no expert is required to prepare the debriefing. Third, several useful features remain unimplemented due to a lack of resources, even though being identified by both workshop participants and developers. We highly recommend that these features be developed in the future. They include enabling notes and logging markers during the debriefing session. Moreover, better options for data synchronization and replay functions are required, particularly when working with video and eye-tracking data. Additionally, export functions and portability would be beneficial. Finally, it would be helpful to support a wider range of data types and to pre-process them, and to incorporate preset training evaluation criteria.

Although it was not initially intended for this use case, participants during the workshop suggested creating user profiles and a collaborative learning platform. This approach could improve training for a specific user group. Taking this concept further, an intelligent learning tool could be developed. In future iterations, the system could provide not only post-simulation feedback but also adaptive interventions during training, addressing both individual and collective competencies and skills.

To conclude, our user-centered design approach can be applied to adapt the web-based tool for different user groups (e.g., human factors experts, instructors, trainees, etc.) and various use cases (e.g., air traffic control, manned/unmanned aviation, etc.). The core concept remains to provide a flexible framework that enables end users to choose the content and features of the web-based tool according to their specific needs. With its adaptable design, the current post-flight debriefing tool enables individuals with IT knowledge to easily understand and tailor it for specific user requirements. In this study, the debriefing tool was customized to meet the specific needs of maritime drone crews and was first implemented in a human factors study within the framework of maritime RPAS operations. While it facilitated debriefing after simulation runs, it also highlighted common issues that can occur when the integration of a tool in a simulation environment is not ideal. The lessons learned from the tool demonstration indicate that further work is necessary to address issues related to time synchronization.

REFERENCES

Allen, R., Hession, P. and Kring, E. (2004) 'Usability analysis of a personal digital assistant based data collection tool for the shipboard training environment', *Naval Air Warfare Center Training Systems Division, Orlando*.

Bordogna, J. (2018) *The Sailboat Exercise*. Available at: https://www.nighttrain.co/the-sailboat-exercise-for-team-retreats/.

Bunn, M., Ditch, P. and Fegley, B. D. 'Aircrew Performance Measurement and Proficiency: A Need for Better Operational Capabilities'. *Advances in Human Factors in Training, Education, and Learning Sciences*. Cham: Springer International Publishing, 255–261.

Burns, J. J., Giebenrath, J. and Hession, P. 'Objective-Based Training: Methods, Tools, and Technologies'. *45 th Annual Conference of the International Military Testing Association*: Citeseer, 116.

German Aerospace Center (2024). Available at: https://www.dlr.de/en/fl/research-transfer/projects/marpas-3-maritime-rpas-operation-3.

German Aerospace Center (2025). Available at: https://www.dlr.de/en/fl/about-us/departments/human-factors.

Greschke, D. A. and Cerutti, S. 'Aircrew Mission Training via Distributed Simulation (MTDS)–Development of the Multi-Country Complex Synthetic Environment'. *RTO HFM Symposium on Advanced Technologies for Military Training*, 13–15.

Henquet, M. and Bellucci, T. 'AI-Assisted Debrief: Automated Flight Debriefing Summarization and Competency Assessment'. *Proceedings of the 2nd International Conference on Cognitive Aircraft Systems*.

International Organization for Standardization (2019) *ISO 9241-210:2019 Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems*. Available at: https://www.iso.org/standard/77520.html.

Jensen, R., Nolan, M., Harmon, N. and Caldwell, G. 'Visually based timeline debrief toolset for team training AAR'. *Proceedings of the 28th Interservice/Industry Training, Simulation and Education Conference*.

Lewrick, M., Thommen, J.-P. and Leifer, L. (2020) 'Introduction', *The Design Thinking Life Playbook : Empower Yourself, Embrace Change, and Visualize a Joyful Life*. Newark, UNITED STATES: John Wiley & Sons, Incorporated, pp. 13–23.

Roth, W.-M. (2015) 'Cultural practices and cognition in debriefing: The case of aviation', *Journal of cognitive engineering and decision making,* 9(3), pp. 263–278. DOI: https://doi.org/10.1177/1555343415591395.

Ryffel, C. P., Muehlethaler, C. M., Huber, S. M. and Elfering, A. (2019) 'Eye tracking as a debriefing tool in upset prevention and recovery training (UPRT) for general aviation pilots', *Ergonomics,* 62(2), pp. 319–329. DOI: https://doi.org/10.1080/00140139.2018.1501093.

Schmorrow, D., Nicholson, D., Lackey, S. J., Allen, R. C., Norman, K. and Cohn, J. (2008) 'Virtual Reality in the Training Environment', *Human factors in simulation and training*, pp. 201–230.

Schrepp, M. (2015) 'User experience questionnaire handbook', *All you need to know to apply the UEQ successfully in your project*, pp. 50–52.

Schrepp, M., Hinderks, A. and Thomaschewski, J. (2017) 'Construction of a benchmark for the user experience questionnaire (UEQ)', *International Journal of Interactive Multimedia and Artificial Intelligence,* 4(4), pp. 40–44. DOI: http://dx.doi.org/10.9781/ijimai.2017.445.

Schrepp, M. and Thomaschewski, J. (2019) 'Design and validation of a framework for the creation of user experience questionnaires', *IJIMAI, 5*(7), pp. 88–95. DOI: https://dx.doi.org/10.9781/ijimai.2019.06.006.

Snipes, S. W., Kristi; Ziegler, Melanie; Jernigan, Gregg; Bechtel, Robert 'Intelligent mission debriefing'. *AIAA Computers in Aerospace VII Conference*, Washington, DC, 596–604.

Stackowiak, R. and Kelly, T. (2020) 'Problem Definition', *Design Thinking in Software and AI Projects: Proving Ideas Through Rapid Prototyping*. Berkeley, CA: Apress, pp. 43–67.

Tannenbaum, S. I. and Cerasoli, C. P. (2013) 'Do Team and Individual Debriefs Enhance Performance? A Meta-Analysis', *Human Factors, 55*(1), pp. 231–245. DOI: https://doi.org/10.1177/0018720812448394.

Wiese, E. E., Freeman, J., Salter, W. J., Stelzer, E. M. and Jackson, C. (2008) 'Distributed After-Action Review for Simulation-Based Training', *Human factors in simulation and training*, pp. 287–299.