

DEVELOPMENT OF A HUMAN FACTORS METHODS BLOG FOR AN AUDIENCE OF GAME DEVELOPERS

Nicholas Mudrick, Lawton Pybus, Michelle Taub, and Francis A. Trowbridge
North Carolina State University

We, as human factors practitioners and researchers, often must describe what we study, how we do it, and why it matters. As a part of those efforts, we developed a blog introducing human factors methods and their applications for an audience of game developers. Each method is discussed and illustrated using examples applied to games. We also describe our goals for the blog's style and structure, and ideas for further refinement.

INTRODUCTION

Human factors practitioners of all stripes must be able to convey the value of their work. Whether applying for faculty or industry positions or research grants, we often need to explain to people outside of our discipline what the field of human factors is and why it is important if only to ensure we can continue doing our work. In a larger sense, if we hope to fit more and more of the systems that people use to them for greater efficiency, effectiveness, and satisfaction, the tools of the trade must be disseminated far and wide. As an organization, the Human Factors and Ergonomics Society has sought to meet this need in a number of ways, including a YouTube series, and advice for Twitter and LinkedIn posts. There have been sessions on this theme at the International Meeting, such as the "Human Factors Elevator Speech." These are worthwhile efforts. Individual members should also make it a part of their work to serve as ambassadors of human factors, expanding public understanding of what we do and why we do it.

The game industry is rife with opportunities for human factors methods. As a large, multi-billion dollar industry covering a large variety of systems, controls, and displays, many game development organizations already retain some human factors practitioners on staff. Those companies give themselves a competitive advantage in their ability to enhance user experience. However, human factors methods are not yet commonplace in the industry, so there is a need to communicate to this audience in particular.



Figure 1. Steam purchase screen

In this work, we developed a blog for an audience of game developers to describe human factors methods and how they can be applied to games using examples of task analysis, heuristic evaluation, workload, and situation awareness assessments.

METHODS APPLIED TO GAME DEVELOPMENT

Task Analysis

As a relevant example of applying a human factors method to a gaming context, we conducted a task analysis (Gillan, 2012) of purchasing Steam 4 on a desktop computer through the Steam entertainment platform (see Figure 1), using observational data of a member of the research team performing the task. Steam is a major service for the

purchase and play of games, and a hub for a large community of gamers. After the observational data was collected, we listed the goals and sub-steps necessary to complete the task. We then arranged the goals and sub-steps within this task hierarchically and detailed the error probability, error consequence, and error criticality for each unit of analysis, according to the SHERPA method (Embrey,

1986). See Table 1 for a representation of this hierarchical task analysis.

Our results suggested that Steam's generally cluttered interface may increase attentional demands and the likelihood of a user missing task-relevant signifiers, becoming distracted by

Game developers and companies could use these results to improve the shopping experience for users. While impulse purchases may be good for the company's bottom line, it may not be in line with the company's long-term goals. Short-term profits may not translate into customer retention, for example. Another possible application would be to modify the interface to improve the user's sense of

Table 1. *Hierarchical task analysis of purchasing a Steam game.*

Task Step	Error Probability	Error Consequence	Criticality of Error
1. Locate Steam App on Computer	Low	Need to download the Steam Application	High
2. Enter in account name and password	Low	Need to create account and password	High
3. Scan interface for relevant input	Low	Spend multiple minutes looking for button	Low
4. Click "Games" icon	Low	Miss icon	Low
5. Identify relevant genre option from drop-down menu	Moderate	Re-scan and identify relevant genre for searching	Low
6. Click "action"	Low	Miss icon, have to relick	Low
7. Narrow by related tags	High	Populate options with irrelevant options	Low
8. Click "Single player"	Low	Miss icon, have to relick	Low
9. Wait for interface to populate with relevant options	Low	List doesn't populate, have to refresh	Low
10. Scan relevant options to identify "Fallout 4"	High	Identify another game that looks interesting	Moderate
11. Click the "Fallout 4" Options	High	Miss and scroll through interface	Moderate
12. Enter in birth date to continue	Low	Forget birthday? Improbable.	High
13. Press enter icon	Low	Miss, have to re-click	Low
14. Scroll to locate purchase icon	High	Miss and scroll, potentially get distracted by other games	Low
15. Press "Add to cart icon"	High	Miss and scroll, potentially get distracted by other games	Low
16. Press "purchase for myself icon"	Moderate	Miss and scroll, potentially get distracted by other games	Low
17. Wait for login information to be identified	Low	Wait	Low
18. If error pops up, click "retry" and redo step 17	Low	Wait	High
19. Enter in credit card information (if new user)	Moderate	Enter in wrong credit card information, have to redo step	High
20. Wait for game to download	Low	Internet can cut out	Low
21. Press Library Icon	Low	Miss, have to re-click	Moderate
22. Locate Fallout 4 on left hand menu	Moderate	Click another game, potentially get distracted and play that game	High
23. Locate and press play icon	Low	Miss, have to re-click	Low
24. Play Fallout 4!	Moderate	Game crash, have to reload	High

irrelevant links (e.g. steps 10 and 11). This is in line with user-generated online discussion of problems with the site. Many users complain of having made too many impulse purchases. Our hierarchical task analysis identified one possible cause for this grievance.

control over their purchases, even if the volume does not decrease.

Heuristic Evaluation

We performed a heuristic evaluation of an action shooter game for smart phones called Max Ammo (see Figure 2). Rather than using Nielsen's (1994) set of standard heuristics, which are not

necessarily the best fit for games in which there is often a desirable level of challenge and difficulty, we used an adapted set by Pinelle, Wong, and Stach (2008) specifically designed for use with games. As the basis for this analysis, the research team collected and analyzed qualitative data from online reviews of the game on GameSpot.com. We then rated the game's adherence to the usability heuristics according to the standard 0 to 4 scale: 1, cosmetic problem; 2, minor problem; 3, major problem; and 4, usability catastrophe (Nielsen, 1994). See Table 2 for a summary representation of the evaluation.

Our results generally favored the design of Max Ammo, with a few exceptions. We noted a common complaint in reviews about the sensitivity of the device to user input, a violation of principle 1 ("Consistency of responses to user actions"). There was a more significant violation of principle 9 ("Instructions, training, and help"). The game only provided instructions for the main gameplay environment, with little to none for other areas of gameplay such as equipment set-up. One evaluator noted that the current design of Max Ammo assumed a degree of prior experience from the user.



Figure 2. Max Ammo gameplay

The developers behind Max Ammo could use the results of such an evaluation to find areas for improvement in future designs that may have been previously unnoticed. To address violations of principle 1, it may be necessary to optimize the game's feedback for specific devices, or to overhaul the mechanism by which the game detects user input. Developers may also wish to rethink how the game delivers instruction for novice users. They may want to confirm that the assumption of prior experience in their user population with some observation or interrogation technique. Alternatively, if it aligns with the goals of the game, they may wish to provide more detailed and

comprehensive help.

Workload

We asked participants to play a session of Tetris, a puzzle game (see Figure 3). We assessed their performance in terms of in-game score, level, time, and number of lines cleared. After the game, we assessed their workload using the NASA-TLX (Hart & Staveland, 1988)

Table 2. *Heuristic evaluation of Max Ammo*

Usability Heuristic	Score (0 – 4)	Comments
1. Consistency of responses to user actions	2 – minor problem	The system does not always register inputs
2. User customization	1 – cosmetic problem	Can only customize audio preferences
3. Predictable and reasonable behavior from computer units	0 – not an issue	Game is very consistent
4. Provide unobstructed views that are appropriate for the user’s current actions	0 – not an issue	Game allows for customizable view options
5. Allow players to skip non-playable and repeated content	0 – not an issue	The storyline can be skipped
6. Provide intuitive and customizable input mappings	0 – not an issue	Novel game – input features only for touchscreen devices
7. Easy to manage controls, with an appropriate level of sensitivity and responsiveness	0 – not an issue	Very easy to use
8. Provide users with information on game status	0 – not an issue	Several options to monitor/check game status
9. Instructions, training, and help	3 – major problem	Only instructions for gameplay are provided
10. Easily interpretable visual representations	1 – cosmetic problem	Should apply if longer missions added to the game

Table 3. *Workload and performance of Tetris players.*

Participant	TLX Score	Score	Tetris		
			Time	Level	Lines
1	80.73	7020	04:17	3	23
2	67.21	7372	07:30	3	28
3	72.40	25298	07:18	6	55

to better understand not only the relationship of workload and gameplay performance, but also the players' satisfaction and comfort with the game (Rubio, Diaz, Martin, & Puente, 2004).

Our results showed the general and expected relationship between workload and performance. That is, participants who demonstrated better performance generally also reported lower workload. However, upon a closer examination of the individual weighted subscales, there were some qualitative differences in the experience of workload between low and high performing players. While there was little variability across players in physical workload (low for all participants) and temporal workload (high), high performing players reported greater mental workload and less frustration. See Table 3 for summaries of results.

These results suggest that high performing Tetris players have adopted a strategy that more efficiently distributes cognitive workload than low performing players, and demonstrate how a workload analysis could be valuable method for game developers who want to better understand the

divide between expert players or "power users" and novice or casual players. A design team may wish to further investigate the underlying causes of these differences with finer grain analyses better suited to answer these questions, such as a cognitive task analysis (Klein, 1995) or assessments of naturalistic decision making (Klein, 2008).

Situation Awareness

We described how a research team could assess the situation awareness (SA) of a user playing a cooperative first-person shooter, *Destiny* (see Figure 4), using the Situation Awareness Global Assessment Technique (SAGAT; Endsley, 1988). To give the reader a good idea of how the method would be performed, we conducted a simplistic hierarchical task analysis of defeating an in-game "boss." We also described how a SAGAT is performed, and generated some example prompts specific to the task designed to assess the three levels of SA: "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" (Endsley, 1995). We discuss some points to consider when conducting a SA assessment. For example, these assessments may be invasive and lack external validity. Furthermore, in the context of a game, it may be desirable for players to lack some degree of SA. We concluded with some discussion of alternative assessment methods.

BLOG STYLE AND STRUCTURE

The goal of the blog was to convey the utility of human factors method in a game

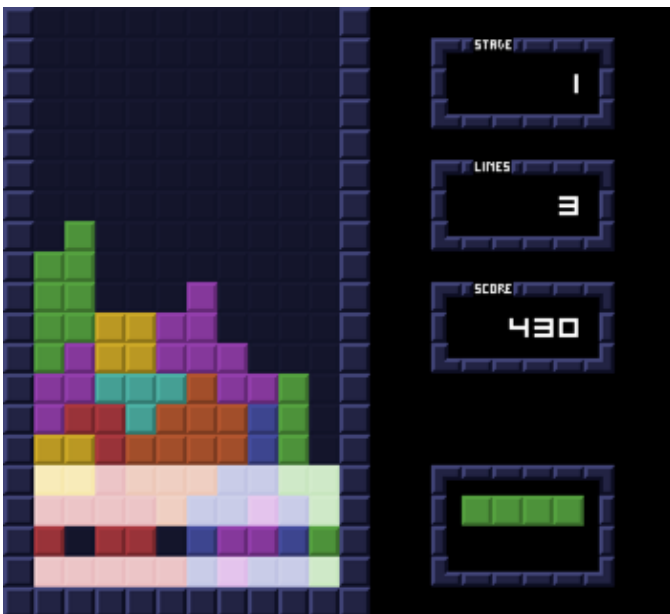


Figure 3. Tetris gameplay



Figure 4. Destiny gameplay

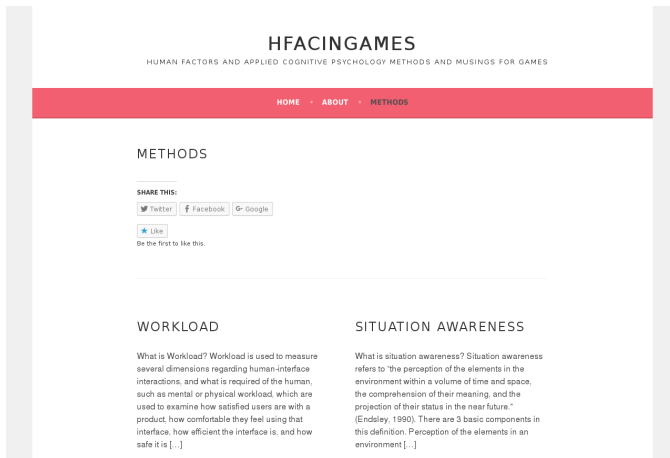


Figure 3. *Human Factors & Applied Cognition in Games* blog

development context, especially to people who may have very limited concepts of what those methods are and can be used to do. We structured the entire blog in an attempt to meet that specific goal. For example, rather than organizing the blog chronologically, as is common with personal blogs, we gave it a content-oriented structure, so that unfamiliar visitors would be able to quickly find relevant information on the specific method they want to learn about.

The individual blog posts on each method were also organized around this goal. Each post begins with a description of the method or the construct it is designed to assess in general. We then provided an illustrative example of the method being applied to a gaming context. Posts concluded with some discussion of how these methods could be applied to improving the game or games in general. Where possible, we included considerations for using the methods appropriately, in order to serve the goals of the game design team and the users who play it. We also point to additional resources, such as tools for conducting the assessments (e.g. the online NASA-TLX; Sharek, 2011) and primary sources.

FUTURE WORK

The blog, titled *Human Factors & Applied Cognition in Games*, is hosted on the Wordpress web publishing service and can be accessed at <http://hfacinggames.wordpress.com/> (see Figure 5). We intend to continue development of the blog with more examples of methods and applications from

human factors practitioners already working in the field. We would also like to refine and improve the structure and style of the blog based on the feedback of members of our target audience. We hope that in time, it will become a point of introduction to our field for the unfamiliar, and a useful resource for developers.

ACKNOWLEDGMENTS

We thank Dr. Douglas J. Gillan for his discussion and training of human factors methods, and for providing useful comments on this project.

REFERENCES

- Embrey, D. E. (1986). SHERPA: A systematic human error reduction and prediction approach. In *Proceedings of the International Topical Meeting on Advances in Human Factors in Nuclear Power Systems*.
- Endsley, M. R. (1988, May). Situation awareness global assessment technique (SAGAT). In *Aerospace and Electronics Conference, 1988. NAECON 1988., Proceedings of the IEEE 1988 National* (pp. 789-795). IEEE.
- Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1), 65-84.
- Gillan, D. J. (2012). Five questions concerning task analysis. In M. A. Wilson, W. Bennett Jr., S. G. Gibson, & G. M. Alliger (Eds.), *The Handbook of Work Analysis: Methods, Systems, Applications and Science of Work Measurement in Organizations* (pp. 201-213). New York, NY: Routledge.
- Hart, S.G., & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139-183). Amsterdam: Holland.
- Klein, G. (2008). Naturalistic decision making. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 50(3), 456-456.
- Klein, G. (1995). The value added by cognitive task analysis. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 39(9), 530-530.
- Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. In B. Adelson, S. Dumais, & J. S. Olson (Eds.), *Proceedings of the ACM CHI Conference of Human Factors in Computing Systems* (pp. 152-158). Boston, MA: ACM Press.
- Pinelle, D., Wong, N., & Stach, T. (2008). Heuristic evaluation for games: usability principles for video game design. In M. Burnett, M. F. Costabile, T. Catarci, B. de Ruyter, D. Tan, M. Czerwinski, & A. Lund (Eds.), *Proceedings of the 25th Annual CHI Conference on Human Factors in Computing Systems* (pp. 1453-1462). Florence, Italy: ACM Press.
- Rubio, S., Diaz, E., Martin, J., & Puente, J. M. (2004). Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and Workload Profile Methods. *Applied Psychology: An International Review*, 53, 61-86.
- Sharek, D. (2011). A useable, online NASA-TLX tool. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 55(1), 1375-1379.