

HCI Usability Testing Methods for the Game Demiurge

COMP210 - Research Journal

1604281

2018-10-10

1 Introduction

Human computer interaction is an integral part of most software, including video games, as the users' successful interaction with the system is a key goal. Because of this, it is important to ensure that systems are intuitive and easy to use, allowing new users to interact without problems. Usability testing is therefore a vital part of software development to ensure the system usability is sufficient.

This journal will examine several methods of usability testing, discussing their benefits and costs, as well as comparing their appropriateness to be used to research the usability of the Demiurge video game combat system.

2 Game

Demiurge is a third person bullet-hell game, in which the player must dodge projectiles or reflect them with their shield, while simultaneously attempting to aim their own projectiles to eliminate enemies. The game is a non-commercial prototype produced by

a team of students. As the game is centred around the combat mechanics, the developers want to analyse the usability of the combat system, specifically how intuitive it is for the player and if the controls cause any confusion.

3 Methods

3.1 Think-Aloud

Think aloud testing is a basic but powerful form of usability testing where the users interact with the system while narrating their thoughts out loud for the observer to record. [1] This gives deep insight into the users' personal way of interacting with the system and quickly highlights areas of confusion or frustration.

There are two main methods of carrying out a Think-Aloud test, concurrent and retrospective. [2] In a concurrent think-aloud, the user relays their thoughts in real time as they use the system. This has the advantage of being spontaneous and linked directly to the users' actions, however users may find it strange to perform in this manner. A retrospective has the user relate their thought process after using the system, possibly alongside a recording of them using the system. Although this may be a more comfortable method for the user, there is a chance that they will forget their exact thoughts, or spend more time thinking about what to say, thereby causing the test to be less genuine. [3] Despite this, a retrospective think-aloud may allow users to more clearly relate complex problems they encountered with the interface, and explain their thought process in more depth . [4]

This method is also cheap and easy to perform, requiring only an observer, the users and the system, which can make it a good starting point to eliminate problems before considering more complex tests. This could be especially useful in the case of Demiurge, as the small team would be able to perform the tests without needing extensive resources. It would also give information on what parts of the combat the players were most

interested in, which could help in balancing other aspects of the combat system, such as pacing or visual feedback.

However, without training or extensive research it can be difficult to analyse the results of a Think-Aloud - especially without a second method such as eye testing to compare with - as user thoughts can be vague and uninformative. This would necessitate an experienced usability tester to perform the tests, or to provide assistance or training to the team. [5]

3.2 Eye-tracking

For more technical data about the system, eye-tracking tests can be conducted to discover how the user interacts with the interface at a deeper level. This method can help observers to understand the data from a think-aloud or similar test, by showing observers what the user was looking at while using the system. It is difficult for people to explain their thought process, as they can process thoughts far faster than they can articulate them. Monitoring eye movements gives an insight into the thought process, as people tend to be thinking about what they are looking at. [6] This also helps to prevent the observer from asking unnecessary questions, as they do not have to prompt the user for their thoughts on where they are looking. [3]

Eye-tracking data can also be analysed separately from other usability tests, simply to provide insight into what parts of an interface drew the users attention the most. This can be used to test the effectiveness of specific system elements, such as adverts.

As Demiurge is a fast-paced combat game, the ability to track player eye movements could help the developers understand how players gather information from the game, which could be used to inform UI design (to ensure it does not cover important areas of the screen) and combat elements, such as the range from which projectiles can be shot at the player. It also avoids distracting the player, (which think-aloud testing may do), allowing them to focus on playing the game in a natural manner. [7] However,

eye-tracking tests require specialised equipment, which is both expensive and needs training to use correctly.

3.3 Surveys

Surveys are a versatile method of data collection that can be tailored to suit any part of the system being tested. One of the largest benefits of surveys is their potential for collecting large amounts of data by reaching a larger number of users than other forms of tests can. [8] As they do not require an observer or specialised equipment, they can be used to assess large groups of users, both in a testing environment, but also as continual testing after release, as surveys can be completed remotely by users. [9] This helps to facilitate ongoing usability testing even after development has been completed and the product has been released.

Larger data sets provide greater opportunities to find meaningful trends and patterns that can be used to identify issues that effect a broad spectrum of users. However this can also be a problem, as without sufficiently large datasets, surveys can fall short, providing a fairly limited breadth of data. Surveys do also require careful thought and research to avoid questions that bias or otherwise influence the user. [10]

4 Conclusion

Ideally, for a more complete review of Demiurge’s combat system, several different tests would be employed so that the weaknesses of certain test types could be mitigated by the other tests. However, due to the small team size and lack of budget, it is far more appropriate to use only one test. While eye-tracking would be useful in analysing the intuitiveness of the combat system’s visual feedback, and help to avoid interface elements that distract or confuse the user, the relatively high cost of the equipment, training and software makes this an impractical approach. [11] [12]

In comparison, a retrospective think-aloud test would not only be cheap, but also

particularly useful for gaining insight into users' strategies and tactics in the game. A concurrent think-aloud may also be useful, but may be fragmented due to the fast-paced nature of the game demanding the users' attention. [13] Although a survey would be an even cheaper method required less training than a think-aloud, the limited scope of the survey may make it more difficult to understand the complex interactions between users and the game, as the users are not able to articulate subtle problems and solutions they may have encountered.

Overall, a retrospective think-aloud would be the most appropriate method of HCI usability testing for the Demiurge game, allowing the developers to understand the ways in which the users play the game, and what they focus on.

References

- [1] J. Nielsen, *Usability Engineering*. Morgan Kaufmann, 1994.
- [2] S. McDonald, H. M. Edwards, and T. Zhao, "Exploring think-alouds in usability testing: An international survey," *IEEE Transactions on Professional Communication*, vol. 55, no. 1, pp. 2–19, 2012.
- [3] B. Freeman, "Triggered think-aloud protocol: using eye tracking to improve usability test moderation," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2011, pp. 1171–1174.
- [4] M. J. Van Den Haak, M. D. T. De Jong, and P. J. Schellens, "Retrospective vs. concurrent think-aloud protocols: testing the usability of an online library catalogue." *Behaviour & Information Technology*, vol. 22, no. 5, p. 339, 2003.
- [5] L. Cooke, "Assessing concurrent think-aloud protocol as a usability test method: A technical communication approach," *IEEE Transactions on Professional Communication*, vol. 53, no. 3, pp. 202–215, Sept 2010.

- [6] —, “Eye tracking: How it works and how it relates to usability.” *Technical Communication*, vol. 52, no. 4, pp. 456 – 463, 2005.
- [7] M. Pretorius, H. Gelderblom, and B. Chimbo, “Using eye tracking to compare how adults and children learn to use an unfamiliar computer game,” in *Proceedings of the 2010 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists*. ACM, 2010, pp. 275–283.
- [8] I. Alam, S. Khusro, A. Rauf, and Q. Zaman, “Conducting surveys and data collection: From traditional to mobile and sms-based surveys,” *Pakistan Journal of Statistics and Operation Research*, vol. 10, no. 2, pp. 169–187, 2014.
- [9] W. Albert and T. Tullis, *Measuring the user experience: collecting, analyzing, and presenting usability metrics*. Newnes, 2013.
- [10] G. Albaum, J. Wiley, C. Roster, and S. M. Smith, “Visiting item non-responses in internet survey data collection,” *International Journal of Market Research*, vol. 53, no. 5, pp. 687–703, 2011.
- [11] X. Yuan, M. Guo, F. Ren, and F. Peng, “Usability analysis of online bank login interface based on eye tracking experiment,” *Sensors & Transducers*, vol. 165, no. 2, p. 203, 2014.
- [12] K. Moran, “Quantitative user-research methodologies: An overview,” 2018. [Online]. Available: <https://www.nngroup.com/articles/quantitative-user-research-methods/>
- [13] S. McDonald, T. Zhao, and H. M. Edwards, “Dual verbal elicitation: the complementary use of concurrent and retrospective reporting within a usability test,” *International Journal of Human-Computer Interaction*, vol. 29, no. 10, pp. 647–660, 2013.