EasyChair Preprint

№ 5233

Designing the BrainTagger Researcher Platform to Automate Development of Customized Cognitive Games

Bella Zhang and Mark Chignell

March 29, 2021

Designing the BrainTagger Researcher Platform to Automate Development of Customized Cognitive Games

Bella (Yigong) Zhang[1] and Mark Chignell[1]

[1]University of Toronto, Toronto, Canada
yigong.zhang@mail.utoronto.ca,chignell@mie.utoronto.ca

Abstract. Serious games have grown significantly in popularity, but proving their scientific validity through research studies is a common hurdle for researchers and game developers. To scale up capacity to collaborate with different groups of researchers, Centivizer Inc. (a University of Toronto spinoff company) has employed a user-centered design process to design a BrainTagger Research Platform (BRP) that will largely automate the development process of its customizable serious games for cognitive assessment. This new development will increase the capacity to gather research data needed to improve game mechanisms and demonstrate game validity.

Keywords: Human Factors Engineering, Serious Games, Cognitive Assessment, User Interface Design, Usability Evaluation.

1 Introduction

The serious games market has grown rapidly since around 2010. It has wide applications in healthcare, education, aerospace, government, retail, media, and entertainment industries. Serious games are defined as games that are created and used beyond entertainment purposes [1]. Compared with traditional cognitive interventions, serious games provide a sense of positivity and playfulness, which can be a good motivational factor when engaging with patients in clinical settings [2]. Cognitive testing, in particular, has been reported to be stressful and cause anxiety, which may impact the accuracy of assessments [3]. Serious games offer a way to carry out cognitive assessment more easily with the gameplay, not only making assessment more fun for the person being assessed but also introducing randomization elements that reduce the possibility of learning effects. This creates opportunities for using the same test repeatedly on the same person without losing validity.

Cognitive tests through serious games have four key advantages over those done in a paper-and-pencil format. Firstly, when a person is engaged with a cognitive assessment in a serious game format, the sense of "being tested" is less apparent. Paper-and-pencil tests can cause anxiety and be stressful, which may impact the accuracy of assessments [3]. Secondly, automated data collection done by software applications is more efficient, effective, and reliable than human transcribed data. This makes it possible to track detailed cognitive changes over time, allowing the game to serve as a health monitoring tool [4]. Thirdly, the use of software makes it

possible for individuals to access cognitive testing in remote and rural areas and allowing health systems to be more proactive in implementing interventions where they are needed. Lastly, the labor costs of paper-and-pencil testing are greatly reduced.

1.1 Problem

Centivizer Inc. (www.centivizer.com) has developed BrainTagger, a suite of 8 (as of this writing) Target Acquisition Games for Measurement and Evaluation (TAG-ME), to assess a person's cognitive abilities. Braintagger began as a design concept [5], which was then implemented as a game for screening for delirium in emergency patients [6] and has seen a succession of games being added to the suite and validated (e.g., [7]). In 2019, Centivizer Inc. made the BrainTagger product freely available to different research groups for research use. As more and more researchers requested using our games in their research studies, a capacity bottleneck was reached, and thus the software development process became repetitive and hard to manage. This bottleneck led to increasing problems with labor costs and inefficiency. There was also a higher probability of human errors as administrators and developers had to manage multiple versions of the games simultaneously and manually. To create a long-term solution to this challenge, we reimplemented the games in the REACT framework to make the software more maintainable, and we created the Braintagger Researcher Platform (BRP) so that researchers could customize games for their research studies without having to have coding skills.

The development of online experimental platforms is still at an early stage and is fast evolving. At the time of this writing, the social distancing requirements of the COVID-19 pandemic have led to many in-person research projects being postponed, canceled, or changed to online experiments. The deployment of online experimental tools is beneficial, and not only in pandemics, because online delivery makes it possible for researchers to continue conducting research online with remote participants, and not just in times of pandemic. A significant advantage of running experiments online is that more diverse groups of participants can be run, including those in remote geographical locations and those with physical and other disabilities. In the long term, the trend of employing online experimental tools in social and behavioral sciences will likely continue to grow, benefiting from the rapid technological advances in big data, software development, and machine learning.

2 Objective

This research project's primary goal is to generate a usable and effective high-fidelity BRP prototype that will serve to scale up the usage of our proprietary games in different research studies. It will eventually contribute to the success of our game design, and we plan to make the games intelligent enough to diagnose brain abnormalities and detect cognitive decline among older adults. This is also an online behavior experiment tool that can be used not only for older adults but also for people of different age groups. The games can be applied for behavioral experiments of other

purposes, such as tracking cognitive development in children, assessing distraction, or detecting symptoms of autism. These games have a high potential to advance future cognitive assessment so that it is more reliable, cheap, fun, and ubiquitous.

3 Method

As BRP is a new product, we approached the design process with a minimal viable product (MVP) mindset. We were not seeking to address all the possible user needs; instead, this prototype aims to visualize a product with the highest customer-valued features by engaging potential users throughout the design cycle. This design project followed the iterative cycle of the human-centered design process: Understand, Create, and Evaluate [8]. In the Understand stage, we first analyzed the game requirements for eight different ongoing research studies and conducted a requirements analysis to understand the common game parameters that needed to be customized (based on the underlying researchers' needs) through emails and interviews. Then we developed personas for three user groups (principal investigators, graduate research students, and undergraduate students) and envisioned common usage scenarios for the new platform. A series of paper wireframes (Fig. 1), mockups, and Figma (https://www.figma.com) prototypes (Fig. 2 - Fig. 4) were created and iterated based on ongoing user feedback in the Create stage. In the Evaluate stage, two rounds of high-fidelity prototype usability evaluation were conducted remotely over videoconferences: 5 participants in the 1st round evaluation (October 2020) and 13 participants in the 2nd round evaluation (January 2021). The prototype was redesigned at the end of each round of usability evaluation based on the findings. The two rounds of usability evaluation were similar in format, each comprising three stages: pre-study questionnaire, scenario walkthrough, and post-study questionnaire with a focus on gathering quantitative usability feedback that included system usability scale (SUS) evaluative ratings.

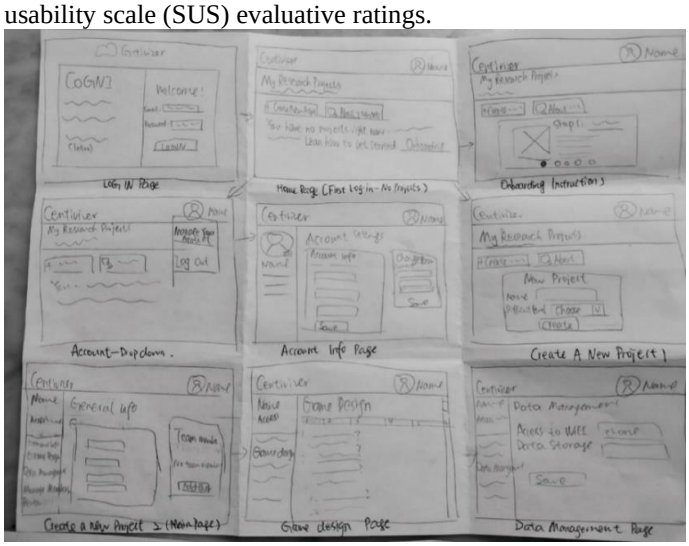

Fig. 1. A sample of the initial paper wireframes

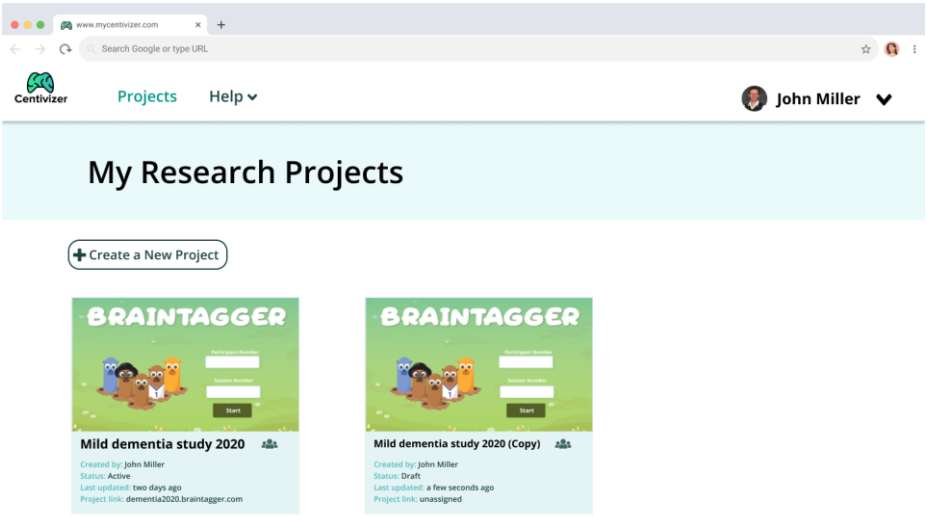

Fig. 2. Screenshot of BrainTagger Research Platform- Project Page

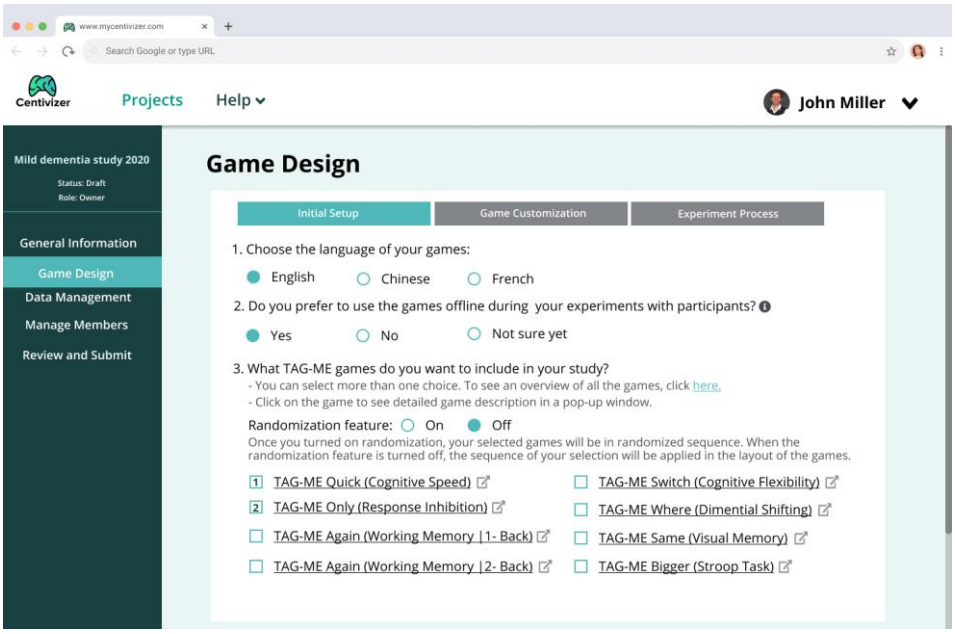

Fig. 3. Screenshot of BrainTagger Research Platform- Game Design Page

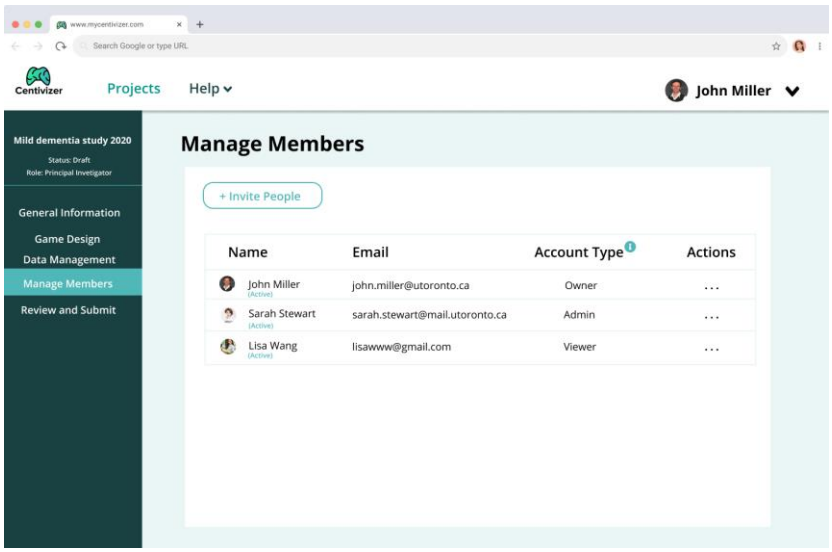

Fig. 4. Screenshot of BrainTagger Research Platform-Management Member Page

4 Results

In both rounds of usability evaluation, all participants commented favorably on the design. When asked about how well the design met the needs as a researcher, the average rating increased from 4.2/5 in the 1st round testing to 4.92/5 in the 2nd round (where 5 is excellent and 1 was poor). All researchers agreed that the researcher platform added high potential value to the TAG-ME games. They described the prototype as "very effective" and "very clean and thoughtful" and expressed their willingness to use it in their future studies. One participant said, "This was my first time viewing the prototype, and I think it is absolutely amazing. It looks very close to being ready for implementation, and I think it is an awesome tool for researchers. I cannot wait to use it!".

The quantitative findings also demonstrated that the prototypes' usability and effectiveness improved greatly between the two rounds of testing. The average SUS scores from both rounds of usability testing were well above the mean of 68 that was obtained over a large number of studies [9]. The average SUS score of the prototype increase from 82.5/100 in the 1st round testing to 87.3/100 in the 2nd round testing. The rating for almost all statements within the SUS scale improved between the two rounds of testing, i.e., the ratings for positive statements increased (Fig. 5) while the ratings for negative statements decreased (Fig. 6).

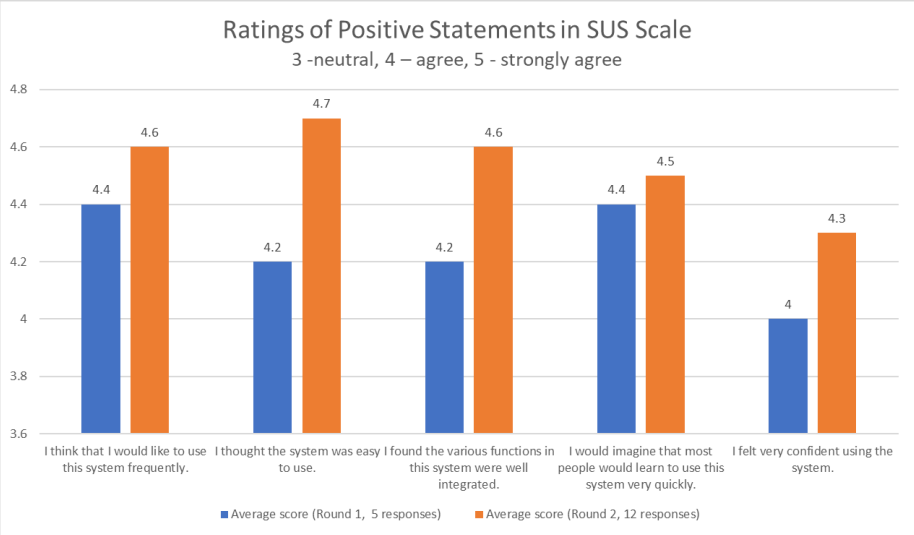

Fig. 5. Rating of positive SUS statements in two rounds of usability testing

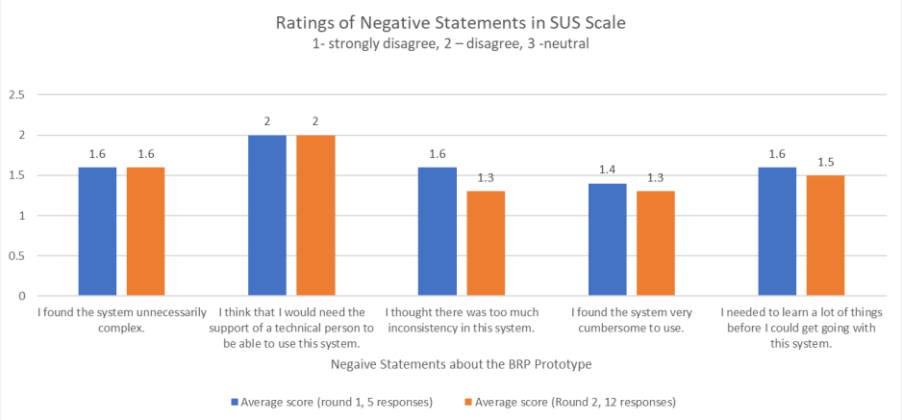

Fig. 6. Rating of negative SUS statements in two rounds of usability testing

5 Discussion

The design of the BRP was successful based on the favorable evaluation feedback received, and the resulting design has been handed to the software team as a specification for development. When completed, the BRP will automate the software development process for customizing those features of most interest to researchers. It will also include collaborative functionalities to facilitate teamwork and communication. As a result, the BrainTagger team will be freed up to work on more complex, special requests from researchers. Meanwhile, the researchers will be empowered to build projects on their own and will have to spend less time

communicating their requirements and waiting for the BrainTagger team to respond to their requests.

We learned from the study that the researchers desire to have more autonomy and flexibility in designing and customizing the online experimental tool without resorting to the development team. Thus, it is important to allow them to make decisions based on the information provided on the interface without having to contact the service provider. It is also important for the user interface to be more informative and explain the functionality and how to make choices wherever possible. As researchers to provide them with the power to generate the experimental tool in real-time. Researchers also highly valued the system's ability to generate recommended parameters based on their target participant group and desired assessment difficulty. Data storage security and sharing agreements are also important concerns when reaching an agreement between the service providers and researchers. Researchers also desire opportunities to collaborate with other researchers working in relevant fields and learn from other people's experiment setups. In our final design, we have incorporated the option to make the project setup public and a page to view the public projects.

In future versions, we will consider increasing the customization ability of the games. For the first version of BPR, the design only allows researchers to choose each game once and, when chosen to prefill, all game parameters are generated with the same level of difficulty. One participant expressed a desire to have: "The option to repeat the same game more than once (for example, to have participants to complete the TAG-Me Only task twice - once with "easier" settings and once with "harder" settings)." An advanced customization feature we could include in the future is to allow researchers to customize each game with a different level of difficulty within their research project.

6 Conclusion

To establish a new, credible research tool requires rigorous efforts to build scientific evidence. The ability to automate software development creates possibilities to promote effective collaboration and scale up capacity to meet researchers' needs. Continuous user engagement coupled with the MVP mindset helped identify and prioritize the most valued features in the product's first iteration. In this product development cycle, we have chosen the waterfall model. The design was finalized before software implementation because there was a lack of software resources during this design process.

This is an ongoing project that aims to not only support ubiquitous cognitive assessment for researchers and others but also to gamify experiments in general, in an approach that we refer to as "Gamified Psychometrics." People who would like to use BrainTagger in their research should contact Mark Chignell (chignell@mie.utoronto.ca).

References

[1] W. Mestadi, K. Nafil, R. Touahni, and R. Messoussi, "An Assessment of Serious Games Technology: Toward an Architecture for Serious Games Design," *Int. J. Comput. Games Technol.*, vol. 2018, 2018.

[2] P. H. P. H. Robert *et al.*, "Recommendations for the use of Serious Games in people with Alzheimer's Disease, related disorders and frailty," *Front. Aging Neurosci.*, vol. 6, no. MAR, p. 54, Mar. 2014.

[3] J. C. Cassady and R. E. Johnson, "Cognitive test anxiety and academic performance," *Contemp. Educ. Psychol.*, 2002.

[4] A. Wilkinson, T. Tong, A. Zare, M. Kanik, and M. Chignell, "Monitoring Health Status in Long Term Care Through the Use of Ambient Technologies and Serious Games," *IEEE J. Biomed. Heal. Informatics*, vol. 22, no. 6, pp. 1807–1813, 2018.

[5] T. Tong, M. Chignell, P. Lam, M. C. Tierney, and J. Lee, "Designing Serious Games for Cognitive Assessment of the Elderly," *Proc. Int. Symp. Hum. Factors Ergon. Heal. Care*, 2014.

[6] T. Tong, M. Chignell, M. C. Tierney, and J. Lee, "A Serious Game for Clinical Assessment of Cognitive Status: Validation Study," *JMIR Serious Games*, vol. 4, no. 1, p. e7, May 2016.

[7] T. Tong, M. Chignell, and C. A. DeGuzman, "Using a serious game to measure executive functioning: response inhibition ability," *Appl. Neuropsychol.*, 2019.

[8] J. D. Lee, C. D. Wickens, Y. Liu, and L. N. Boyle., *Designing for people: An introduction to human factors engineering*. CreateSpace, 2017.

[9] J. R. Lewis and J. Sauro, "Item Benchmarks for the System Usability Scale," *J. Usability Stud.*, 2018.