Towards First Person Gamer Modeling and The Problem with Game Classification in User Studies

Katherine Tarre  
Florida International University  
ktarr007@fiu.edu

Adam S. Williams  
Colorado State University  
AdamWil@colostate.edu

Naphtali D. Rishe  
Florida International University  
ndr@acm.org

Armando B. Barreto  
Florida International University  
Barretoa@fiu.edu

Francisco R. Ortega  
Colorado State University  
fortega@colostate.edu

ABSTRACT
Understanding gaming expertise is important in user studies. We present a study comprised of 60 participants playing a First Person Shooter Game (Counter-Strike: Global Offensive). This study provides results related to a keyboard model used to determine an objective measurement of gamers’ skill. We also show that there is no correlation between frequency questionnaires and user skill.

CCS CONCEPTS
• Human-centered computing → User models;

KEYWORDS
User Modeling, User Studies

ACM Reference Format:

1 INTRODUCTION
User evaluation of interfaces is a common-practice in 3D User Interfaces (3DUI)—whether it is a Virtual Reality (VR) study or a traditional device comparison experiment. Understanding the type of user for an interaction provides for a deeper analysis of the results. For example, if a game controller (GamePad) is used and the user is experienced with that GamePad (e.g., Microsoft Xbox One controller), it is important to separate them from inexperienced user in the analysis. The current classification approach has been ad-hoc, using non-standardized questionnaires to determine participant gaming experience [5]. Some studies have touched upon finding measurements for game levels (see [4, 9]).

We present a user study to understand how to categorize gamers in first-person shooter (FPS) environments. Our study investigated if typical questionnaires (e.g., [5]) to separate users for statistical analysis can be significant. We also looked at creating a FP (first person) model to improve categorization, relating in game performance with self-described skill. We are providing all data upon request for further studies. We provide a user study comparing a GamePad vs a Keyboard+Mouse in order to understand user skills with 60 subjects. We demonstrate that questionnaires (e.g., [5]) do not provide reliable classification data. We provide an early attempt of a keyboard+mouse model to classify FPS gamers.

2 CURRENT LITERATURE
Game modeling has been studied from the point of view of the user either with their behavior or psycho-physiological signals [6]. Understanding the skill of the player is important for classification (e.g., expert gamer). For example, Huang et al. ran two studies that analyzed gameplay in order to understand how skill relates to practice and habit [3]. In their analysis of the First Person Shooter (FPS) game Halo: Reach, they found that people who practiced moderately by spacing out their gameplay gained more skill per match than those who played intensively [3]. When analyzing the strategy game Starcraft II, researchers found that the main difference between skilled and non-skilled players was their use of hotkeys (customized keyboard shortcuts that execute commands quickly). One study showed that people who preferred FPS games had faster reaction times than those who played Multiplayer Online Battle Arena (MOBA) gamers. FPS gamers also showed a reduced ability to cancel prepotent motor responses than MOBA or massively multiplayer online role-playing game (MMORPG) gamers and made more errors after stop signals than those who preferred MOBA and MMORPG. Furthermore, FPS gamers had lower inhibitory control than MOBA gamers [2]. Additional Studies, such as [1, 7, 8, 10] discuss gamer’s behavior and skills to a great extent.

3 EXPERIMENT DESIGN
Our study consisted of 60 participants, 70% men and 30% women. Ages ranged from 19 to 38 with a mean of 24.4 years of age and median of 23.5 years of age. In our sample, 68% of the participants had experience with PC games, 78% had experience with console games, and 38% had experience with hand-held consoles. The majority of our participants, 85%, had experience with PlayStation consoles, while only 47% had experience with the Xbox consoles. 95% of participants had experience with FPS games, out of which 43% of them had experience with some version of Counter-Strike.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

VRST’18, November 28-December 1, 2018, Tokyo, Japan
© 2018 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-6086-9/18/11
https://doi.org/https://doi.org/10.1145/3281505.3281590
The order in which the keyboard+mouse or gamepad (see Figure 1) were used were randomized to prevent bias. Participants were asked questions about past gaming and input device experience. The questions are adapted from [5]. Additionally, gameplay recordings and the locations of shots (head, body, arm, etc.) were collected. Overwolf was used to record keyboard and mouse inputs. Participants played 7 rounds with both the keyboard and the mouse for a total of 14 rounds. The average time in game for participants was 12.5 minutes.

The equipment used during the experiment ran on a Alienware Laptop with an Intel core i7 6820HK clocked at 2.70 Ghz (4-core w/ HT), 16GB RAM, DirectX 12, Windows 10 64-bit, a GPU Nvidia GeForce GTX 980M and a resolution of 1920x1080. The server, which hosted part of the game, ran on an Alienware Desktop model X51 R3 with an Intel core i5 6400HK clocked at 2.70 Ghz (2-core w/ HT), Nvidia GeForce GTX 960, 16GB RAM, and Windows 10 64-bit. Counter Strike (CS:GO) run version 1.35.4.

4 ANALYSIS

For the purpose of the study, the primary measures of our data were: damage given to the bots and damage received by the video game character. A Shapiro-Wilk test found no significant deviations from normality in the data of interest. The device that was first used had no significant impact on the scores achieved for both damage given and damage received. This was the case for both the GamePad and the keyboard and mouse ($t(29)=-2.04, p<.05$). Participants did consistently better when using the keyboard and mouse in comparison to the GamePad ($t(59)=-4.65, p<.001$). This was the case when comparing each paired group as well as all of the participants at once. No correlations were found between self-reported player skill and player performance in game.

Next we considered the possibility of developing a model to accurately predict the skill level of participants in video games. First, we considered the results from the GamePad, however, scores with this device tended to be more erratic and a number of participants had trouble with the commands so we could not develop a model based on its results.

For the Keyboard and Mouse, we found that the frequency in which keys are selected was useful. With this information, we were able to develop a model for the self-rated skill level of participants on a 5-point scale, 0 to 1. Our model takes into account the the frequency at which participants press the W and S keys during the course of one full game – as defined in our experiment design – and whether or not the participant has experience with PC games, 0 for no experience and 1 for some experience. The final model is shown in Equation 1. The model shows all participants started at a baseline of about 0.25 (a little skilled), with the score decreasing when the W (forward) key is pressed repeatedly but increases when the S (back) key is pressed. Experience with PC games ($\text{EP}_{\text{PC game}}$) also increases the overall skill level by 0.3, which is about one step in our 5-point scale. This makes sense because in our observations, we noted that experienced players tend to move continuously and fall back when facing the bots, whereas inexperienced gamers would move step-by-step or in waves and have slower reaction times. All variables in this model are significant at a 1% level of significance. We have no concerns with the residuals or the normality assumption.

\[
\hat{y} = 0.2674 - 0.0014 \times W_{\text{ke}} + 0.0046 \times S_{\text{ke}} + 0.304 \times \text{EP}_{\text{PC game}} \tag{1}
\]

5 CONCLUSION

It is important to understand the limitations of this study while taking the results under consideration. One of our most important findings is that the use of ad-hoc questionnaires to determine users skill level cannot be considered reliable at this point. While they may provide a way to understand users, there is no correlation with the skill displayed by the gamer. Nonetheless, it was possible to find a relationship between user skill in game, when using the keyboard and mouse, and the skill level they consider themselves to have. Our model (see Equation 1) still requires further validation.

It is important to understand the type of gamers participants are during a user study. We have shown that questionnaires to classify users are not reliable and that classifying users by skill assessment is always better than self classification or experimenter classification. Future work includes a skill classification for GamePads and further validation for our keyboard+mouse model.

REFERENCES


