Including Non-Gamers: A Case Study Comparing Touch and Motion Input in a 3D Game for Research

Isabelle Kniestedt, Elizabeth Camilleri, and Marcello A. Gómez Maureira

University of Malta, Msida, Malta
ikniestedt@gmail.com, bethcamilleri94@gmail.com,
ma.gomezmaureira@gmail.com

Abstract. While digital games are becoming increasingly popular as a choice for research stimuli, their complex nature brings about challenges. The design of the games and designers' reliance on established conventions may hinder their use in research, particularly with 'non-gaming' test subjects. In this study, we explored how players performed using a 1-to-1 motion control scheme using a tablet's gyroscope to control the camera as compared to a traditional touch-based joystick in a 3D first-person game. Results showed that players – particularly those less experienced with games – found the game more enjoyable and exciting with motion controls than with joystick controls. Additionally, while experienced players performed better than inexperienced ones when using the joystick, this difference was not present when using the motion controls. We therefore believe motion-based control schemes can be beneficial in making research using games more accessible to a wider range of participants, and to limit influence of prior gaming experience on gathered data.

Keywords: Digital games, input controls, research stimulus

1 Introduction

Digital games, both commercial and custom made, have a long-standing history of being used as research tools in a variety of fields and studies [5, 20]. Some advantages of using games as research stimuli are their ability to make abstract experiment tasks more approachable and understandable [12]. Due to their design they can also provide ideal tools to elicit certain emotions [32], and they provide a safe, virtual environment to explore research topics through experiments that might otherwise be dangerous, deemed unethical [11], or physically impossible for participants to take part in. Using games in this context comes with certain challenges. Commercial, off-the-shelf games have the potential to introduce uncontrolled or unknown variables into the experiment setup that can influence test results. Games developed specifically for research can limit the amount of variables in play, but budget restrictions do not always allow for dedicated designers and developers, leading to the use of very straightforward gaming tasks that often do not offer the same affordances as professionally developed games [11], therefore not taking full advantage of the medium.

Another concern is accessibility to test participants, an often varied and undefined 'target audience'. Games are more ubiquitous than ever before, with the mobile industry projected to take up almost a quarter of the entire games market in 2017 [3]. Casual games are characterized by being especially easy to pick-up-and-play by those who would not usually consider themselves ‘gamers’ [2]. When using more complex game systems, a certain level of gaming experience is generally required (e.g. being able to use an analog game controller, or knowledge of using the ‘WASD’ keys on a keyboard to move around in a game environment). In order to use games to effectively test human behavior, however, participants cannot be limited to ‘gamers’. Ideally, games that are to be used for such test experiments should therefore be easily accessible to a general population and prior gaming experience should not influence the data gathered in such experiments.

The case study presented here was motivated by this design concern of accessibility while developing a tablet game for research purposes. A requirement of this project was that the player explores a 3-dimensional environment, for which we developed a motion-based control scheme utilizing the strength of natural mapped interactions [19]. These controls use the tablet’s gyroscope to directly control the first-person in-game camera, resulting in the player looking around the virtual environment as if they were taking a picture with the tablet in the real world. This was then compared to ‘traditional’ controls using an on-screen joystick operated by touch input, as is often seen in commercial mobile games [9].

With this comparison we tried to answer two questions. The first was whether it would be less challenging for inexperienced players to navigate a virtual 3D environment using the motion controls, as compared to the touch-based equivalent of an analog game controller. The second question was whether the motion controls would introduce any variables in the experiment setup that researchers should be aware of if they utilize this input method in their own studies. This paper describes the testing of the game using a mixed-method approach to data gathering in an attempt to capture various possible effects of the control scheme on player experience. The resulting discussion is informed by our experience in iteratively developing and testing the game. Our aim is to contribute to the growing body of work related to the use of video games as research stimuli by suggesting that careful design of input methods can aid in issues regarding participant selection and the quality of data, an approach that so far, to our knowledge, has been overlooked.

2 Background

Using games as research tools is not a new notion; experiment tasks assigned to participants in which their performance is scored can often be labeled or perceived as a game [5]. Games, however, have become increasingly varied, complex, and sophisticated over time, as has the technology that mediates them. This can make them suitable for new types of research [29], but also makes them less predictable and harder to utilize in a controlled experiment setting. With many studies across disciplines using game-like stimuli in some capacity [1, 13, 29], the body of work specifically concerned with using games in this capacity has also grown. So far, efforts are widespread and varied, ranging from specific contributions focused on a particular experiment task or field [7, 8, 23] to more general guidelines concerning stimulus design and experiment setup [5, 6, 11, 12, 20, 26]. Notable recurring topics are the advantages and disadvantages of using commercial and custom-made games, how to select appropriate games for a specific study, approaches to data logging, and how to work towards being able to generalize results across studies.

Participant selection is mentioned as part of these studies, but solutions are lacking, with the consensus seemingly being that extensive logging and strategic participant selection are the best way to approach the issue of varying player ability. McMahan et al. [20] describe the importance of considering player background when using commercial games as a research stimulus, stating that prior game experience can affect enjoyment and anxiety, as well as a participant's attitude and motivation, when participating in a research study using digital games. They suggest to either include or exclude participants based on their experience, or to passively gather data on participant background to take prior experience into account during data analysis. This coincides with suggestions from other authors, stating the importance of detailed logging to explain differences in data between participants or participant groups [5]. Similarly, Järvelä et al. [11] suggest being "selective" with participants and to pay special attention to prior gaming experience. They state that basic skills in playing digital games are preferred, as time spent on learning takes time away from the experiment tasks, and a lack of basic skills is likely to negatively influence quality of the data.

This conclusion is far from ideal, as it means excluding possible participants based on a non-essential skill not necessarily related to the topic of a given study. It also raises a secondary concern, namely that a certain amount of practice time with the game is needed to gain consistent results across participants, or for the research task to 'stabilize' [13]. Even when limiting participants based on prior experience, not all games are the same, and different participants will require different times to become accustomed to the game they need to play. Any effort to ease this process, and include participants of different backgrounds, should be worth exploring. With gaming becoming increasingly common as a pastime, it would seem solutions could be found within the game industry.

Researchers have paid specific attention to the usability of mobile games and the effects of input methods on performance, for instance by comparing virtual and physical controls [4]. In these comparisons, virtual controls often perform worse than their physical counterparts, likely due to a lack of tactile feedback. As Teather and MacKenzie [26] state, there is relatively little research comparing gyroscope-based controls (or 'tilt' controls) to touch controls (using an on-screen equivalent of a traditional game controller or direct touch input through gestures). In their research they found that previous studies yielded conflicting results in terms of player performance when being subjected to the different controls. Their own results show little difference in performance between the two styles (tilt-based motion and touch control), although touch control still slightly outperformed tilt in the highest game levels. It is important to note that their study, as well as others exploring similar topics, all used 2-dimensional games as their experiment stimulus, with none exploring the types of controls in a 3D environment. One exception is a comparison of three first-person shooter games on *iPod Touch* [9], where one of the games offered a tilt-based input method to control the camera combined with touch input using virtual joysticks for movement. Hynninen [9] concludes that tilt input performed worse than pure touch-based controls using virtual joysticks and swiping gestures. None of the input methods performed well, however, when compared to traditional mouse and keyboard controls in similar games, suggesting it was not just the tilt controls that impacted performance, but rather a poor adaptation of controls familiar to first-person shooter players to a new platform. It is also important to note that while testing 3-dimensional games, the game used in their study does not directly map the position and rotation of the device to that of the game camera, but rather uses a tilting motion to gradually turn the camera, which likely lacks the speed and precision necessary for the fast game-play associated with first-person shooters.

Popularized by Nintendo when they released the Nintendo *Wii*, the game industry has shown a trend for producing different types of controls in an effort to broaden their audience to include casual gamers and those who would not consider themselves 'gamers' at all. Referred to as 'natural interaction techniques' [19], these types of controls mimic realistic movements, e.g. swinging the *Wii Mote* like a racket while playing virtual tennis. Since players already know how to perform the intended action, the learning curve is lower than when learning an abstract control system. Based on earlier research showing that natural interactions can provide greater usability than non-natural interactions for some tasks, MacMahan et al. [19] explored the effects of natural interactions in *Mario Kart* [22] on the Nintendo *Wii*. Their findings show players performing worse when using the motion controls than when using the physical game controller, with suggested reasons for the bad performance being the game's poor use of the Wii remote's sensors, latency issues, and the use of large muscle groups over smaller muscles which contribute to a lack in precision. This is indicative of a similar issues found when comparing touch-based controls with tilt input, as described above. Combating this issue of precision, as well as making the controls as intuitive as possible, was one of our main concerns while developing the game, which we will describe in detail next.

3 The Game

The game used in this study is a single-player tablet game that was originally created for cognitive science research into human foraging behavior [27]. We were asked to create a game based on a 2-dimensional visual searching task, which would be used to test the same behavior in a 3D environment. A central design concern was making the game accessible to non-gamers, as the underlying cognitive task is not exclusive to those who play games frequently. In our experience of performing research with games we found that 3D navigation can be a significant obstacle for novice gamers. Additionally, participants that do not consider themselves gamers tend to be concerned with performing well during an experiment using a game, which could impact data gathered (especially related to performance and emotional state) [20].

In the game, the player takes on the role of a squirrel gathering food for its family. The player controls the squirrel from a first-person perspective and explores a park where 'target' (positive) and 'distractor' (negative) objects are spread across the environment. The default mode for the game has the player on a two-minute timer in which they try to gather as many of the target objects as possible. Objects are collected upon collision, meaning players simply run into the objects to collect them. Collecting tar-

Fig. 1. Game with joystick controls (left) and early play-tester using motion controls (right)

get objects grants points, while distractor objects temporarily restrain the player and negatively impact their score by resetting it to zero.

We relied on fairly simple game mechanics, using points and a timer reminiscent of older ‘arcade’ style video games. In order to emphasize the exploratory aspects of the game, we attempted to create an interesting game-play experience by putting players in an unfamiliar perspective: that of a small animal exploring a large environment. This was done through the scale and design of the digital environment, the positioning of the camera and its field of view, movement speed and sound effects.

The controls developed for this game, in this paper referred to as the **motion** control scheme, use the tablet’s gyroscope to orient the first-person camera to where the player points the device. In this sense it works as if one would take a picture using the tablet’s camera, but instead of looking at real-life surroundings it looks into the virtual environment. In practice, this has the player physically turning their body to turn around in the game, as well as point the tablet up and down to look above or below them in the game environment. This control scheme’s limitation is that it requires the player to either stand up or be seated on a swivel chair in order to play the game comfortably, as being seated on a static chair hinders movement. The only other control in the game is the ‘move’ button, which is located at the bottom left of the screen. By holding the button pressed the player will move forward in the direction they are looking. Releasing the button will make the player stop moving.

In addition to the motion controls, we developed a second input method more akin to controls commonly used in commercial mobile 3D games. We refer to this scheme as **joystick** controls. The joystick controls have the same movement button in the bottom left of the screen. However, instead of positioning and turning the tablet to look around, players use a virtual joystick on the bottom right of the screen, mimicking a standard video game controller and control schemes found in commercial games.

3.1 Iterations and User Testing

Development and refining of the game took place over a 13 month period, in which both the game and the controls went through multiple iterations. The controls were subject to user testing from several weeks into the project and continued to be tested and developed throughout the time leading up to the experiment described in this study. We conducted a pilot study to the work presented here with 12 participants after the completion of the first fully functional prototype. Both control schemes were tested and compared, and the results informed both the design of this experiment and adjustments to the game's parameters (e.g. movement speed, camera field of view, and joystick responsiveness). Overall responses to the game were positive, leading us to assume that the majority of participants would find the game engaging and therefore ensuring that the results of the study would not be impacted by unsatisfactory game-play. Further testing was done in the context of cognitive science research, which led to additional adjustments to the design and helped fine-tune both control schemes. Through this iterative process, we feel confident that both control methods are comparable to those of commercial games, and should not negatively impact our findings.

4 Experiment Design

We tested two experiment conditions: **playing the game using the motion controls,** and **playing the game using the joystick controls.** Participants played both conditions in succession, playing two 'rounds' of the game, each with a fixed length of two minutes, under each condition. Starting conditions were alternated between participants to control for potential influences due to experimentation order — players with even participant numbers started with motion controls, while players with odd participant numbers started with joystick.

For the experiment, the target and distractor objects in the environment were walnuts and acorns of two different sizes; players were asked to collect the small nuts (0.8 scale from the default size), while the large nuts (1.2 scale from default) acted as distractors. The environment, as well as the position and type of objects found within, was consistent between tests. This made the control scheme the only difference in the game between conditions.

After having played both conditions (for a total of four game rounds) the participants answered a questionnaire in which they ranked the two conditions with respect to several emotion states. During play we also recorded psycho-physiological measurements using a wristband sensor as an additional, unbiased indicator of player experience [15], and logged game metrics to analyze player behavior and performance.

4.1 Measurements

The survey consisted of two parts: questions related to demographic information, and preference ranking of the two conditions. Demographic data gathered included age and gender. We also documented gaming frequency, differentiating between frequency in playing 'mobile' and 'other' games (i.e. PC or console) as a general separation of 'casual' and 'hardcore' gamers [14]. Players could choose from 5 options: 'Less than a few times a year', 'A few times a year', 'A few times a month', 'A few times a week', and 'Multiple times a week'.

Players were then asked to state their preference between the two conditions with respect to several emotion states: *enjoyment*, *challenge*, *distraction*, *frustration*, and *excitement*. Each question followed the same structure (e.g. 'Which of the two control schemes do you feel was more enjoyable?'). We based our decision to ask players to annotate their experience in the form of pairwise preferences rather than rely solely on ratings (e.g. Likert scale) on research which shows that people rate emotions better in relative terms than in absolutes, and thus yield more reliable annotations of player experience [21, 31]. Through a 4-alternative forced-choice (4-AFC) protocol, players could choose the first or second condition that they played, as well as indicate no preference by selecting either 'Both' or 'Neither'. We used this 4-AFC protocol so as to not force a preference where there potentially was none, with the intent of making player choices clearly motivated and more meaningful [30]. Below each question was a comment box where participants could freely elaborate on their choice. In addition to these questions, players were asked to rate their general enjoyment of the game (irrespective of the control scheme) on a scale from 1 to 10.

To measure player behavior and performance we logged game metrics, including amount and type of items collected (positive and negative), player position (recorded at 10 Hz), and camera rotation (10 Hz). Each of these were logged with a time code so that player route and collection events could be checked against sensor data. One of the advantages of using a digital stimulus for research is the ability to record player behavior and analyze it afterwards [5]. In this case we were particularly interested in the way and amount that participants moved in the virtual environment between test conditions, as well as their performance in regards to item collection.

Psycho-physiological data (also referred to as biometric data) was collected using the *Empatica E4* wristband sensor [18, 24], which tracked *heart-rate* (reported at 1 Hz, based on blood volume pulse captured via a PPG sensor) and *electro dermal activity* (EDA) (4 Hz). Over the last decade, metrics and biometrics have become a part of the quality assurance practice of several notable game development studios. Similarly, game research has adopted the use of biometrics to study a variety of topics, e.g. game features, events, and emotional effects [15]. While different studies have shown conflicting results and the connection between psycho-physiological data and emotional states does not yet show itself to be completely reliable, we included biometric measurements in our study to gain an additional indicator of how the different controls might affect the players' experience and to provide an additional data point on a subject that is still much debated [10, 15]. Additionally, we deemed it important to add measurements that provided unbiased data to counter possible designer-bias.

4.2 Procedure

Tests were carried out over the duration of a week at the University of Malta. Participants were gathered through a combination of convenience and purposive sampling. As sampling from the university campus showed a pattern of less gaming experience, we purposely included participants from the university's digital games course.

The game was played on an *iPad 4th Gen*. The laptop used to gather data from the iPad and biometric sensor after each session was set up on a central table, with chairs for the researchers on one end and the swivel-chair for the participant on the other. Participants were asked to read through a general information sheet and sign the consent form, while the researchers assigned them an ID and prepared the game and sensor. Next, they asked the participants to put on the sensor, or helped them to put it on if requested.

Two baseline readings of the biometric data were gathered against which the measurements taken during the game were compared: 30 seconds with the participant sitting still (used as a baseline to the data collected while playing with the traditional controls), and 30 seconds while moving the arm with the sensor in circles (used as a baseline to the data collected while playing with motion controls). During the baseline readings, participants were asked to count along with a metronome ticking at 80 beats per minute with the goal of still engaging the participants in a task (as they would be in the game). Based on previous tests we found that while using the joystick, players were largely motionless. While playing with the motion controls, players move their arms up and down and turn around in the chair. It should be noted that in none of our tests this led to participants moving very fast or suddenly; their movements were generally controlled and steady. The baseline readings should therefore be comparable to the play sessions, and the motion of the game should not have negatively influenced the sensor data.

Participants played two rounds (of two minutes each) with the first control scheme (alternating which one they started with between participants). Scores from collected objects aggregated between rounds, and players were encouraged to aim for a high-score at the beginning of the experiment session. Once the timer ran out, players saw their score and were able to progress independently to the next round. After finishing the second round they were returned to the game's main menu. One of the researchers then changed the control scheme, before handing them the iPad a second time. They then played another two rounds with the other control scheme, before being asked to take off the sensor. The session concluded with the participant filling in the survey.

For both game metrics and biometric data, we only looked at the second round that participants played in each condition. This was so that players had one round to get used to each of the controls, and limit the effects of players getting used to a control scheme on the data.

4.3 Data Processing

Data from the metrics logged in-game and the psycho-physiological sensor was processed before evaluation. First, game logs were parsed for positions and rotation of the player camera for each participant, in order to analyze the in-game behavior of participants. Similarly, data was parsed for positive and negative pickup events as well, which were used as a measure of player performance. We aggregated the spatial distance between subsequent player positions and the difference between subsequent camera rotations to explore differences in overall player movement.

Five different sub-measures were derived from the sensor data: *Median*, *Median Absolute Deviation (MAD)*, *Slope*, *Travel*, and *Onsets* (for EDA only). Additionally, each sub-measure was calculated for each of the baseline readings (i.e. motionless and in motion) taken at the start of the experiment. We then divided the sub-measures of an experiment session by the sub-measures of the respective baseline to derive a value that expresses the relative change of a participant's sub-measure as compared to their baseline. An exception was the sub-measure of Slope, which is expressed in measure per minute. Outliers in the data (values larger than $3 * MAD$ [16]) were removed and the sensor data was pre-processed with a Gaussian filter (window width = 4 * measure frequency).

After calculating the sub-measures across all participants we performed a Shapiro-Wilk test of normality under an alpha level of 0.05, and again removed outliers (values larger than $3 * MAD$, this time across participants rather than across sensor measures over time for each individual participant) that resulted in $p < 0.05$. Sub-measures that still showed a deviation from normality after this were omitted from further analysis and are not reported in this study. All of the data processing was facilitated by custom-made scripts, while statistical analysis of measures across all participants was conducted with the statistical software JASP [17].

5 Results

We concluded the research experiment after having tested $N = 31$ participants, 48.4% of which were female. The median age was 20.50 ($M = 23.57$, $SD = 8.25$). The median for the frequency of playing mobile video games was 3 out of 5 ($M = 2.94$, $SD = 1.44$), and 3 out of 5 for non-mobile video games ($M = 3.03$, $SD = 1.35$). To recall, this measure was taken through a nominal scale going from 'Less than a few times a year' (converted to 1), 'A few times a year' (2), 'A few times a month' (3), 'A few times a week' (4), to 'Multiple times a week' (converted to 5).

In order to determine whether individual measures differed between the two test conditions, we performed Bayesian Paired Samples T-Tests, opting for a Cauchy prior width of 0.707, in accordance with research by Wagenmakers et al. [28]. The value of the Bayes Factor BF (BF_{10}) indicates the likeliness that a given hypothesis is not equal to its null-hypothesis, i.e. the assumption that there is no significant difference between the tested conditions. A BF_{10} value of 1 indicates that there is an equal chance of the hypothesis being different from the null-hypothesis as there is of them being similar. A value lower than 1 indicates that the null-hypothesis is more likely. Unlike classical hypothesis testing, the Bayesian T-Test can therefore be used to confirm the null-hypothesis, rather than only reject it [25], making it particularly helpful for this study. The results for each individual measure, as well as their corresponding mean and standard deviation, are shown in Table 1. Mean values for game metrics correspond directly to player actions (distance corresponding to meters, rotations corresponding to degrees), while mean values for sensor measures are percentages with regards to their respective baselines (a mean of −33 indicating a decrease of 33% over baseline). The majority of measures indicate evidence *for* the null-hypothesis. In Table 1 this is indicated by values in **bold** for $BF_{10} < 0.333$, which means that for a given measure it is at least 3 times more likely that empirical (small) differences are *not* significant. Three measures suggest anecdotal evidence *for* significant differences (indicated by (†) for $BF_{10} > 1$, but $BF_{10} < 3$). For these measures we further conducted the more widely used Student T-Test, which showed significant differences ($p < 0.05$) for two results: 'Camera Rotations' – $t(30) = 2.237$, $p = 0.033$, and 'Heart-rate Median' – $t(29) = 2.224$, $p = 0.034$.

Measure	**Mean**$_{Motion}$	**SD**$_{Motion}$	**Mean**$_{Joystick}$	**SD**$_{Joystick}$	**BF**$_{10}$	**error%**
Distance Traveled	115.274	23.698	106.554	24.852	0.752	9.89e-5
Camera Rotations	4557.140	1387.413	3842.211	1074.532	† 1.655	2.92e-8
Total Pickups	18.000	8.896	17.774	8.578	**0.193**	3.36e-4
Positive Pickups	17.226	9.694	17.290	9.093	**0.192**	3.42e-4
Negative Pickups	0.774	1.117	0.484	0.851	0.533	5.04e-5
Heart-rate Median	1.173	11.414	-6.303	19.137	† 1.284	7.19e-8
Heart-rate Slope	-0.613	3.799	-0.651	6.287	**0.207**	1.36e-4
Heart-rate Travel	-28.120	55.118	-33.883	51.636	**0.276**	2.31e-4
EDA Median	33.263	33.133	58.145	67.534	† 1.059	2.02e-7
EDA MAD	119.205	204.899	116.577	196.896	**0.201**	4.04e-5
EDA Travel	-17.911	64.016	-0.649	75.966	**0.269**	1.35e-4

Table 1. Bayesian Paired Samples T-Test of game metrics and sensor measures. Measures with (†) indicate *anecdotal evidence for* a significant difference between conditions ($BF_{10} > 1$, but $BF_{10} < 3$). Measures in **bold** indicate at least *moderate evidence against* a significant difference between conditions ($BF_{10} < 0.333$).

After playing the game in both test conditions, participants were asked to rate their enjoyment between 1 ('did not enjoy at all') and 10 ('enjoyed it a lot'), which resulted in a median rating of 7 out of 10 ($M = 7.19$, $SD = 1.2$), with 87.1% stating they would play again. Participants were further asked to rank which of the two conditions was more *enjoyable*, *challenging*, *distracting*, *frustrating*, and *exciting*. The distribution of the responses is shown in Figure 2. Clear preference was given to motion controls in both enjoyment (77.4%) and excitement (80.6%). Joystick controls were more frequently deemed challenging (48.4% — Bayesian Binomial Test, against > 0.25: $BF_{10} = 44.531$) and frustrating (51.6% — $BF_{10} = 44.531$).

Participants were asked to elaborate on their survey choices through written comments. Out of 31 participants, 10 did not comment on their choices, while the rest of the participants filled out all or some of the comment fields. For the states of *enjoyment* and *excitement*, half of participants made comments related to moving around being more fun than remaining stationary. Positive aspects mentioned were turning speed, ease of motion, and a higher level of "immersion". Four people specifically commented on preferring the joystick controls, finding movement easier without physically turning around and one participant reporting slight motion sickness. Among the reasons for the joystick being considered more *challenging* was having to divide attention between two actions (i.e. turning and moving). Similar arguments were given for *distracting* and *frustrating*, with divided attention (e.g. through keeping track of thumb position) being listed as a reason for the joystick to distract from the game.

Finally, we looked at potential influences that could impact the comparison between the two experiment conditions. Chi-Squared Tests between 'Play Frequency' (mobile and non-mobile) and self-reported preferences did not show evidence for preferences being influenced by how frequent participants play video games (all tests $p > 0.05$). Looking at correlations (Pearson's r, $p < 0.05$), we found that 'Play Frequency (non-mobile)' correlated with 'Distance Traveled (Joystick)' ($r = 0.404$, $p = 0.024$). A

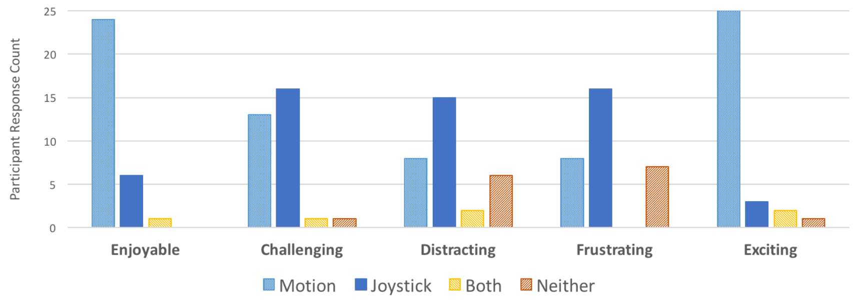

Fig. 2. Frequency distribution of self-reported preferences regarding which condition was more 'Enjoyable', 'Challenging', etc. Participants could forgo ranking by choosing 'Both' or 'Neither'.

Bayesian Independent Samples T-Test (Cauchy Prior 0.707) between start conditions showed that the majority of measures had a $BF_{10} < 0.333$ and none had a value above $BF_{10} > 1$, indicating that no significant differences were due to starting condition.

6 Discussion

During this study we were interested in effects on player experience and performance, as well as the ease of use of the motion as compared to joystick controls, between participants with varying levels of game playing experience. Regarding player performance, measured by the amount of items collected, we did not see a significant difference between conditions. This suggests that regardless of prior gaming experience, participants performed equally well in both test conditions. This is in contrast with other studies, where motion-based controls negatively impacted performance [9, 19]. We consider this a positive result for our study, as it suggests that the motion controls do not cause players to under-perform like other tilt-based input methods or when compared to joystick controls used in commercial games. We believe the direct mapping of the camera to the position and rotation of the tablet to be the main reason for this difference, as it provides a higher level of accuracy and control than other tilt-based input methods.

Other measures of particular interest were distance traveled and degrees of camera rotation. With the goal of the game being the exploration of an environment and the collection of items dispersed throughout said environment, we consider the amount that participants traveled indicative of how easy it was to move around. Frequency of playing non-mobile games (i.e. PC or console games) showed a positive correlation to distance traveled in the joystick condition. It stands to reason that those who are more likely to be familiar with analog controllers would need less time to master the joystick input scheme. This suggests a learning curve for those who did not have this prior experience. The fact that we did not find significant evidence of a difference in the motion control scheme suggests that the learning curve was equal regardless of prior experience. With the survey results suggesting that participants in general enjoyed using the motion controls more, we consider this the strongest indication that the motion controls were able to bridge the gap between participants of varying gaming backgrounds and experience.

A student t-test showed a significant difference in camera rotations between conditions, with an increased amount of rotations when using motion controls. This is not completely surprising, as every movement of the tablet is recorded by the game, while the use of the joystick requires purposeful action to turn the camera. However, the difference is large enough that we do not think accidental movement is the sole explanation. The physical movement of the motion controls requires more effort from the player to turn and look around. A possible explanation for this difference could be, as participants stated in their written and verbal comments, that the act of moving physically was more fun than using the joystick and caused them to feel more ‘in the game’, comparing it to a virtual reality experience. Alternatively, it is also a possibility that the joystick controls were generally perceived as less interesting to use, which led to participants being more efficient in how they used them. Overall, it would suggest that the motion controls contributed to the sense of exploration and ‘looking around’ we attempted to stimulate in our original design.

Overall, there are not many indications of the motion controls causing subconscious responses in players’ biometric measures, suggesting the control scheme can be used in research without introducing hidden variables that could influence data collection of this kind. We did find evidence in the heart rate median, which was higher compared to the relative baseline when using motion controls. Although we did not find significant correlations between game ranking and heart rate data, we speculate that this difference can be explained for similar reasons as the increase in camera rotation. Since the motion controls were deemed more exciting by the majority of players, the increase in the heart rate can be another expression of the participants’ excitement while playing. Here we want to re-iterate the importance of recording a baseline reading that takes the motion of the controls into account in order to accurately compare the biometric data. As described earlier, we attempted to take a baseline measure that reflected the amount of motion the players would experience while playing the game for each condition. We therefore believe the difference can be explained by the participants’ emotional state, rather than through the act of being in motion. However, we make room for the possibility that, despite our efforts, the movement of the controls still had some effect on the data, which could be remedied by changes in the baseline procedure.

We recognize that our sample size is relatively small, but the data so far seems to indicate that the motion controls can indeed make research games more accessible to inexperienced participants. Tests throughout the game’s development have shown that the motion controls are easier to get used to and require less explanation for non-gamers. This does not mean, however, that these particular controls will work for every study. First, there is the requirement for participants to move. While not a concern for our purposes, space and other limitations (e.g. using more sensitive sensor equipment) could prevent motion controls from being used in a particular study. Second, this game was developed as a visual searching task. Moving the tablet to search for items is closer to the physical action of foraging than the use of a game controller, making the controls an appropriate choice. While research participants were encouraged to get a high score, (turning) speed was not an essential part of the game experience (as it would be in, for example, a first-person shooter game). Accuracy was a necessity as target and distractor items could be positioned close together, requiring careful navigation to only collect the correct one. As it stands, we consider the current control scheme to be useful for games that require the navigation and exploration of 3D environments where either no or only simple actions are required. A single button, or similar input, at minimum is required to move the player forward, and an additional button could be implemented for simple interactions (e.g. talk, jump, open a door, read a message, pick up an object). Careful interface design would be necessary to make the use of an additional button intuitive (e.g. 'move' and 'action' buttons should not be placed too close to one another on the screen, as we experienced in earlier versions of the game).

At this time, we do not exclude any particular game from using this type of control, as appropriate interface and input design can mediate a lot of potential issues with games that require more complex controls. We do, however, foresee potential issues with certain types of games, for instance first-person shooters and other games that require quick reflexes. The current control scheme does provide the accuracy needed for such a game, but a person's physical turning speed could be an issue if the game requires a certain level of speed. Additionally, rapidly turning around could cause physical issues (e.g. dizziness or nausea) in some participants. By designing the game in such a way that it does not require players to turn around in too large increments too fast, these issues could be mediated. However, this type of controls is likely to be less fast and accurate in general than those using a mouse and keyboard, where players often can move in multiple directions (e.g. backwards and by strafing) using only the small motions of button presses. Comparing the motion controls to a setup such as this was out of the scope for this study, but would be valuable to explore in future work. While limiting movement can ease the learning curve for inexperienced players, it would be interesting to see the two compared directly. Another type of game for which it would be harder to utilize this type of control is a racing game, or any other game that through its metaphor (e.g. sitting in a car) restricts a player's movement. Follow-up studies using the control scheme in different games could illuminate which other types of research experiments could benefit from using this type of control specifically.

7 Conclusion

This study explored the effects of two input methods on player experience and performance when interacting with a 3D game in a research setting. The concern regarding prior gaming experience among participants and its influence on data quality has led to the suggestion of excluding participants inexperienced with games from research experiments. It was our goal to bridge the gap between experienced and inexperienced participants, allowing inexperienced players to be a part of research studies without impacting data. We tested two controls: a motion-based control scheme mimicking the use of a tablet's camera, and 'traditional' video game controls in the form of a virtual joystick. Our results show that the motion controls were generally preferred by players, deemed more enjoyable and exciting, and can limit differences in player behavior among participants with varying game playing experience. Players' performances were not negatively impacted by the motion controls, and much of the biometric data points to evidence against significant differences between the two conditions. We therefore suggest that the use of carefully designed, non-conventional controls, such as the motion-based interaction explored in this study, can limit impacts due to prior gaming experience without introducing hidden variables in the experiment setup. We hope this study motivates further exploration into this topic.

8 Acknowledgments

We would like to thank all test participants that took part in both the pre-tests and the final experiment for their time and feedback. We also want to thank Francesca Borg Taylor-East for her contribution to the initial development of the game. Last, but not least, we are very grateful to Prof. Ian Thornton for his continued enthusiasm and support as we developed this game for his research and continued to improve it into a universal tool to be released to the academic community.

References

1. Calvillo-Gámez, E., Gow, J., Cairns, P.: Introduction to special issue: Video games as research instruments. Entertainment Computing 2(1), 1–2 (2011)
2. Casual Games Association: Casual games market report 2007 (2007), retrieved from: http://www.casualgamesassociation.org/pdf/2007_CasualGamesMarketReport.pdf
3. Casual Games Association: Towards the global games market in 2017: A broad look at market growth by screen and region (2014), retrieved from: http://cdn2.hubspot.net/hubfs/700740/Newzoo_Games_Industry_Growth_Towards_2017.pdf?t=1455182596586
4. Chu, K., Wong, C.Y.: Mobile input devices for gaming experience. In: User Science and Engineering (i-USEr), 2011 International Conference on. pp. 83–88. IEEE (2011)
5. Donchin, E.: Video games as research tools: The space fortress game. Behavior Research Methods 27(2), 217–223 (1995)
6. Elson, M., Quandt, T.: Digital games in laboratory experiments: Controlling a complex stimulus through modding. Psychology of Popular Media Culture 5(1), 52 (2016)
7. Harteveld, C., Sutherland, S.C.: Personalized gaming for motivating social and behavioral science participation. In: Proceedings of the 2017 ACM Workshop on Theory-Informed User Modeling for Tailoring and Personalizing Interfaces. pp. 31–38. ACM (2017)
8. Holmgård, C., Togelius, J., Henriksen, L.: Computational intelligence and cognitive performance assessment games. In: Computational Intelligence and Games (CIG), 2016 IEEE Conference on. pp. 1–8. IEEE (2016)
9. Hynninen, T.: First-person shooter controls on touchscreen devices: A heuristic evaluation of three games on the ipod touch. Master Thesis (2012)
10. Järvelä, S.: Measuring digital game experience: Response coherence of psychophysiology and self-reports. Master Thesis (2017)
11. Järvelä, S., Ekman, I., Kivikangas, J.M., Ravaja, N.: Digital games as experiment stimulus. Proceedings of DiGRA Nordic 2012 pp. 6–8 (2012)
12. Järvelä, S., Ekman, I., Kivikangas, J.M., Ravaja, N.: A practical guide to using digital games as an experiment stimulus. Transactions of the Digital Games Research Association 1(2) (2014)

13. Jones, M.B., Kennedy, R.S., Bittner Jr, A.C.: A video game for performance testing. The American Journal of Psychology pp. 143–152 (1981)
14. Juul, J.: A casual revolution: Reinventing video games and their players. MIT press (2010)
15. Kivikangas, J.M., Chanel, G., Cowley, B., Ekman, I., Salminen, M., Järvelä, S., Ravaja, N.: A review of the use of psychophysiological methods in game research. Journal of Gaming & Virtual Worlds 3(3), 181–199 (2011)
16. Leys, C., Ley, C., Klein, O., Bernard, P., Licata, L.: Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. Journal of Experimental Social Psychology 49(4), 764–766 (2013)
17. Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., et al.: Jasp (v.0.7) (2015), Computer software
18. McCarthy, C., Pradhan, N., Redpath, C., Adler, A.: Validation of the empatica e4 wristband. In: Student Conference (ISC), 2016 IEEE EMBS International. pp. 1–4. IEEE (2016)
19. McMahan, R.P., Alon, A.J.D., Lazem, S., Beaton, R.J., Machaj, D., Schaefer, M., Silva, M.G., Leal, A., Hagan, R., Bowman, D.A.: Evaluating natural interaction techniques in video games. In: 3D User Interfaces (3DUI), 2010 IEEE Symposium on. pp. 11–14. IEEE (2010)
20. McMahan, R.P., Ragan, E.D., Leal, A., Beaton, R.J., Bowman, D.A.: Considerations for the use of commercial video games in controlled experiments. Entertainment Computing 2(1), 3–9 (2011)
21. Metallinou, A., Narayanan, S.: Annotation and processing of continuous emotional attributes: Challenges and opportunities. In: 2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG). pp. 1–8 (April 2013)
22. Nintendo: Mario kart wii (2008), Published by Nintendo
23. Raffert, A., Zaharia, M., Griffiths, T.: Optimally designing games for cognitive science research. In: Proceedings of the Cognitive Science Society. vol. 34 (2012)
24. Ragot, M., Martin, N., Em, S., Pallamin, N., Diverrez, J.M.: Emotion recognition using physiological signals: Laboratory vs. wearable sensors. In: International Conference on Applied Human Factors and Ergonomics. pp. 15–22. Springer (2017)
25. Rouder, J.N., Speckman, P.L., Sun, D., Morey, R.D., Iverson, G.: Bayesian t tests for accepting and rejecting the null hypothesis. Psychonomic Bulletin & Review 16(2), 225–237 (April 2009)
26. Teather, R.J., MacKenzie, I.S.: Comparing order of control for tilt and touch games. In: Proceedings of the 2014 Conference on Interactive Entertainment. pp. 1–10. ACM (2014)
27. Thornton, I.M., Kniestedt, I., Camilleri, E., Gòmez Maureira, M., Kristjánsson, Á., Prpic, V.: Simulating foraging in the wild using an ipad. Presented at ECVP 2017 (2017)
28. Wagenmakers, E.J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Selker, R., Gronau, Q.F., Dropmann, D., Boutin, B., et al.: Bayesian inference for psychology. part ii: Example applications with jasp. Psychonomic Bulletin & Review pp. 1–19 (2016)
29. Washburn, D.A.: The games psychologists play (and the data they provide). Behavior Research Methods 35(2), 185–193 (2003)
30. Yannakakis, G.N.: Preference learning for affective modeling. In: 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops. pp. 1–6 (Sept 2009)
31. Yannakakis, G.N., Martínez, H.P.: Ratings are overrated! Frontiers in ICT 2, 13 (2015), `http://journal.frontiersin.org/article/10.3389/fict.2015.00013`
32. Yannakakis, G.N., Paiva, A.: Emotion in games. Handbook on affective computing pp. 459–471 (2014)