Age and sex effects on Super G performance are consistent across internet devices

Andrew Hooyman and Sydney Y. Schaefer

School of Biological and Health Systems Engineering, Arizona State University, Tempe, AZ, USA
{andrew.hooyman, sydney.schaefer}@asu.edu

Abstract
There have been recent advances in the application of online games that assess motor skill acquisition/learning and its relationship to age and biological sex, both of which are associated with dementia risk. While this online motor learning assessment (called Super G), along with other computer-based cognitive tests, was originally developed to be completed on a computer, many people (including older adults) have been shown to access the internet through a mobile device. Thus, to improve the generalizability of our online motor skill learning game, it must not only be compatible with mobile devices but also yield replicable effects of various participant characteristics on performance relative to the computer-based version. It is unknown if age and sex differentially affect game performance as a function of device type (keyboard versus touchscreen control). Thus, the purpose of this study was to investigate if device type modifies the established effects of age and sex on performance. Although there was a main effect of device on performance, this effect did not alter the overall relationship between performance vs. age or sex. This establishes that Super G can now effectively be extended to both computer and mobile platforms to further test for dementia risk factors.

1. Introduction
Current research on dementia is focused on identifying individuals in the preclinical phase through the use of biomarkers [1], [2]. Biomarkers are measurable indicators of a biological state or condition that can be used to diagnose or predict the progression of a disease. In the case of dementia, biomarker development can help identify individuals in the early stages of the condition before symptoms become severe, and at a time when a serious game for dementia care can be most impactful [3]. One potential biomarker related to dementia-specific neurodegeneration is the ability to acquire and retain motor skills. Although previous research has focused on the targeting of cognitive decline in dementia [4], studies have shown that changes in motor skill learning can be an early indicator of cognitive decline in individuals with dementia as well [5]–[7]. Thus, motor skill learning deficits could be employed as an
enrichment strategy in clinical trials for dementia to help identify patients that would benefit the most from participation in the trial. However, for any enrichment strategy to be successful it needs to be generalizable at a population level. This has led to the creation of large web-based patient groups or “cohorts” that can be used to identify individuals in the preclinical phase of dementia [8], [9]. These web-based cohorts allow for the collection of large amounts of data from individuals across the country, making it easier for researchers to identify potential biomarkers and to conduct large-scale clinical trials.

Recent research has shown that motor skill acquisition can be accurately measured through personal computers and individual test kits [10], [11]. However, these methods are currently only available through a single mode of delivery, either computer-based or through a mailed test kit. This is a limitation as it may exclude certain individuals who may not have access to a personal computer or the ability to receive a mailed test kit. Additionally, as of now, neither of these assessments have been adapted for use on mobile devices. This could lead to an unintentional exclusion of certain demographic groups who are more likely to access the internet through mobile devices rather than personal computers, such as Black/African American or Hispanic/Latino [12]. This is a significant issue to address, as these racial/ethnic groups are known to be at much a higher risk of being diagnosed with dementia than Non-Hispanic white individuals [13]. This highlights the importance of developing mobile-compatible assessments to ensure equitable access to tools and resources for measuring motor skill acquisition, and to better understand how these tools can act as dementia biomarkers in the future.

To date, only one study has effectively collected and analyzed motor data from a mobile device in relation to neurodegeneration [14]. However, this study primarily focused on measuring gross motor performance, such as finger tapping, which may not be sensitive enough to detect cognitive and dementia risk factors in the preclinical phase for the purpose of enriching clinical trials as a biomarker. Previous laboratory-based research has shown that more complex motor tasks, such as functional reaching tasks, are better indicators of cognitive status among older adults than simpler tasks, such as measuring maximal grip strength [15]. However, simply converting a lab-based task to a smartphone app does not guarantee reliable results [16], and some physical tasks are simply not translatable to the online space. Furthermore, as highlighted by previous working groups [17], apart from the conversion of a lab-based task to an online platform, the strength and weakness of any serious game is also dependent on the cost and ease of use.

Given these non-trivial challenges in translating computer-based measures to mobile devices, the purpose of this study is to investigate the effect of internet device type (computer or mobile) on motor skill acquisition using a motor-cognitive game called Super G [18], [19]. The study aimed to analyze if there was a difference in skill acquisition when using a computer vs. mobile device, and how this difference relates to the participant's age and sex. Age and sex are important to consider when evaluating the consistency and reliability of performance between device types, as they are known to be significant risk factors for Alzheimer's disease [20], [21] as well as significant determinants of internet use and video game play [22], [23]. Additionally, it is known that phone behavioral characteristics can vary based on age and sex [24]. By examining if the effects of age and sex on Super G performance are different when played on a computer (laptop, desktop) vs. a mobile device (smartphone, tablet), the study aims to broaden the use of digital motor games in understanding motor skill acquisition and its relationship to dementia biomarkers.

2. Methods and Material

This study was approved by the Institutional Review Board of Arizona State University (Study 00015247). Participants were recruited through Amazon Mechanical Turk (MTurk) with the
requirement of a 99% approval rate from Amazon MTurk, residency in the United States, and being 18 years or older. Two batches were run through MTurk, each lasting 7 days, with the goal of recruiting 200 participants for each batch. The first batch recruited participants to play Super G on a computer, while the second batch recruited participants to play on a mobile/touchscreen device. Participants who participated in the first batch were excluded from participating in the second batch.

Participants who clicked on the batch link were taken to an informed consent page. They were informed that they needed to complete 75 trials of the game to receive compensation, which was confirmed by a code they received after completing all trials. The 75-trial requirement is based on prior research that demonstrated that 75 trials is the minimum number of trials necessary to generate participant-specific acquisition curves [19]. After agreeing to participate, participants filled out a survey asking for their MTurk ID, age, sex, race, ethnicity, and education level (less than high school, high school equivalent, some college but no degree, associate’s degree, bachelor’s degree, master’s degree, doctorate or professional degree). After completing the survey, they were given a link to the Super G website. This process was the same for each batch except for the mobile batch, which received a link and a QR code for easier access to the site on a mobile device.

Participants who continued to the Super G website were first given instructions on how to play the game and control the astronaut. These instructions were: “You will play as the astronaut, Super G, whose only goal is to explore new worlds and distant galaxies (16 Total)! If on a PC or desktop, to control Super G you will use the right arrow key to move Super G forward and the left arrow to move Super G backward. If on a mobile or tablet, to control Super G you will touch the right side of the screen to move Super G to the right and touch the left side of the screen to move Super G to the left. Every trial Super G will start on the planet to the left of the screen, the start planet. Super G can only leave the start planet once the blue halo around the start planet disappears. If Super G leaves too early then the trial restarts. If Super G leaves the start planet at the right time then Super G will have 4 seconds to land on the goal planet, located on the right of screen. To successfully land on the goal planet Super G must stay on the goal planet for 1 complete second. If Super G is successful then a reward chime will sound, fireworks will appear and a new planet will appear on the next trial! Try to travel to all 16 planets until you get to the final galaxy!” A screenshot of the website can be viewed in Supplementary Material. To keep track of each participant’s progress, they were asked to create a user account using their MTurk ID. This ensured that the data from the game and survey could be combined for analysis. The Super G website is hosted by a secure third-party service and does not collect any personal information, such as IP addresses. The site was equipped with a secure sockets layer and cookies were disabled to ensure participants’ privacy and security. All participant information was stored in a secure, dual-factor authenticated database with salt-hashed usernames and passwords. The data collected during the game included the astronaut’s position and acceleration, the date and time of each trial, the number of trials completed, the highest score, the total planets landed on, the device’s screen refresh rate, and the method of control (keyboard arrow keys, touchscreen, or idle). The astronaut’s position and acceleration were recorded at 60 Hz.

Participants were instructed that the goal of the game is to help the astronaut, “Super G”, travel to as many planets as possible using either the computer arrow keys or their device’s touchscreen. Each trial was 4.5 seconds in length. The astronaut started on the start planet, surrounded by a blue halo which disappeared 1.5 seconds into the trial. If the astronaut left the bounds of the start planet prior to the halo disappearing, then the astronaut spawned back to the center of the planet and the trial timer restarted. With the three remaining seconds the participant used either the left and right arrow keys of a keyboard (on their computer) or the right of left side of a touchscreen (on their mobile device) to apply positive or negative force to the astronaut. For a trial to be successful, the astronaut must stay within the target planet for
1 complete second. If this occurred, then fireworks blasted off the target planet and a reward tone was made. Then on the next trial the astronaut spawned back at the start planet, which was now rendered as the previous target planet and a new planet was rendered as the target planet. Thus, time in target was the primary dependent variable for each trial. Because participants repetitively practiced the task over the course of 75 trials, motor skill acquisition was operationally defined as the improvement in time in target over time, with longer times indicating better performance up until the 1-second limit. The Super G game was coded in JavaScript through the open-source HTML5 online game engine PlayCanvas. Visual of the Super G game can be seen in Figure 1.

![Super G Game](image)

Figure 1. Visual display of the Super G game. Note astronaut in the start location (left side) and the target location (the red planet, right side).

2.1 Statistical Analysis
To analyze differences in performance between groups based on device type we performed independent t-tests and chi-squared tests across variables of age and sex, as well as education, hour of the day played, race and ethnicity. A p-value of .05 or less would be considered statistically significant.

To analyze the effect of device type on motor skill acquisition as a function of age or sex, we used a linear mixed effects model. With our primary outcome variable as time in target for each trial we modeled two 3-way interactions as our fixed effects. The first 3-way interaction modeled an effect of device by trial by age and the second 3-way interaction modeled an effect of device by trial by sex. For this analysis we modeled trial on a logarithmic scale, with a random intercept for participant and random slope for trial. This is because modeling trial as a logarithm fits Super G better than a linear fit, based on AIC values in prior studies [18] such that time in target generally improves more early on in training and less so later in practice. We included covariates of education, screen refresh rate (in Hz), and the average hour of the
day that participants completed the game (00 – 24). Hour was transformed using the cosine function to better account for the cyclical nature of time, i.e. hours 00 and 24 are closer in proximity that 00 to 12 or 12 to 24. To account for possible structural collinearity, which may lead to inflation of the standard error of the variable estimates, all numeric variables were scaled with a mean of 0 and standard deviation of one: log(Trial), age, screen refresh rate, and average hour of game played. To identify the presence of any variance inflation, we calculated the variance inflation factor of each variable with a criteria that the factor should be below a level of 5. If the factor did exceed 5 we would report this and temper our interpretation of the variable estimate and significance. We did not include education, race, or ethnicity as covariates due to the imbalance within the sample (i.e., primarily white and non-Hispanic with a high bias toward people with a bachelor’s degree), as we were not adequately powered to interpret any result from these analyses across these different variables.

3. Results

3.1 Participant Characteristics

A total of 168 participants participated in the computer group and 111 participants participated in the mobile device group. There were no differences between groups based on age (p > .05) or sex (p > .05). Full demographic breakdown, by race, ethnicity, and education, of participants by group can be seen in Table 1. These initial results demonstrate that there was no bias in terms of study participation based on device type due to participant age, sex, education, hour of the day played, race or ethnicity.

Table 1. Demographic breakdown by age, sex, race, ethnicity, hour of the day game was played, screen refresh rate of the device, and education between groups that performed Super G using the computer keyboard versus a touch screen.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Keyboard</th>
<th>Touch</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>168</td>
<td>111</td>
<td></td>
</tr>
<tr>
<td>age (M/SD)</td>
<td>38.86 (10.49)</td>
<td>38.28 (9.95)</td>
<td>0.64</td>
</tr>
<tr>
<td>sex (M/F)</td>
<td>89/79</td>
<td>54/57</td>
<td>0.56</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>137</td>
<td>88</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>10</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>15</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Mixed</td>
<td>3</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Native American</td>
<td>3</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>13</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Not Hispanic or Latino</td>
<td>153</td>
<td>101</td>
<td></td>
</tr>
<tr>
<td>Hour of the day</td>
<td>14.63(2.98)</td>
<td>14.03(5.13)</td>
<td>0.26</td>
</tr>
<tr>
<td>Device Screen Refresh Rate (Hz)</td>
<td>61.65(25.08)</td>
<td>65.48(17.12)</td>
<td>0.13</td>
</tr>
<tr>
<td>Education</td>
<td>0.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than High School</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>High School Equivalent</td>
<td>25</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Some College but No Degree</td>
<td>27</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>Associate’s Degree</td>
<td>13</td>
<td>16</td>
<td></td>
</tr>
</tbody>
</table>
3.2 No effect of device type on skill acquisition and age

The results of our mixed effects model demonstrated that there was no effect of device type on skill acquisition dependent on age ($\beta_{\text{Device(Touch)}:\text{Trial}:\text{Age}} = .99$, $t(20640) = .14$, $p = .89$). This suggests a participant’s age did not impact their performance on the task due to the type of device they used to complete the task. There was also no device by age interaction ($\beta_{\text{Device(Touch)}:\text{Age}} = -10.15$, $t(271) = -.43$, $p = .67$). This suggests that a participant’s initial performance was also unaffected due to their age and the type of device they used. There was a significant interaction of trial by age ($\beta_{\text{Trial}:\text{Age}} = -10.09$, $t(20640) = -2.33$, $p = .02$), which indicates that older participants acquired the motor skill at a slower rate over the course of training compared to younger participants, which we have demonstrated previously [10]. There was also a significant main effect of age ($\beta_{\text{Age}} = -42.15$, $t(271) = -2.96$, $p = .003$) which indicates that older participants had lower time in targets overall compared to younger participants. Again, as noted above, these age differences were independent of which device type was used. All results for these interactions can be seen on Table 2 and visualized in Figure 2.

![Figure 2. Predicted time in the target planet across trials based on participant age and device used (computer versus mobile). Lighter blue shading of the curve represents older age participants compared to darker blue shading.](image)

3.3 No effect of device type on skill acquisition and sex

Additionally, the mixed effects model also demonstrated that there was no effect of device type on skill acquisition dependent on sex ($\beta_{\text{Device(Touch)}:\text{Trial}:\text{Sex(Male)}} = 15.81$, $t(20640) = 1.12$, $p = .26$). This indicates that a participant’s sex did not impact their skill acquisition due to the type of device they used to complete the task. There was also no interaction between device and sex ($\beta_{\text{Device(Touch)}:\text{Sex(Male)}} = -13.84$, $t(271) = -0.3$, $p = .77$). This suggests that a participant’s sex did not impact their initial performance on the task due to the type of device
they performed the task. There was a significant effect of trial by sex ($\beta_{\text{Trial}:\text{Sex(Male)}} = 40.97$, $t(20640) = 4.63$, $p < .0001$), such that participants who identified as male improved at a faster rate than participants who identified as female. There was also a main effect of sex ($\beta_{\text{Sex(Male)}} = 181.77$, $t(271) = 6.22$, $p < .0001$, estimated Cohen's $D = .78$) which indicates that males had higher initial performance than females. Again, as noted above, these sex differences were independent of which device type was used. All results for these interactions can be seen on Table 2.

Table 2. List of coefficient estimates for each term within the linear mixed effects model in ms. Each numeric term was scaled with a mean of 0 and a standard deviation of 1. *Identifies p-values < .05, ** p-values less than .001, *** p-values less than .0001.

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Estimate (ms)</th>
<th>Standard Error</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>427.01</td>
<td>21.44</td>
<td>19.94</td>
<td>&lt;.0001***</td>
</tr>
<tr>
<td>Device(Touch)</td>
<td>-121.81</td>
<td>33.55</td>
<td>-3.63</td>
<td>0.0003***</td>
</tr>
<tr>
<td>Log(Trial)</td>
<td>117.88</td>
<td>6.45</td>
<td>18.29</td>
<td>&lt;.0001***</td>
</tr>
<tr>
<td>Age</td>
<td>-42.15</td>
<td>14.23</td>
<td>-2.96</td>
<td>0.0033**</td>
</tr>
<tr>
<td>Sex(Male)</td>
<td>181.77</td>
<td>29.21</td>
<td>6.22</td>
<td>&lt;.0001***</td>
</tr>
<tr>
<td>Hz</td>
<td>70.87</td>
<td>11.05</td>
<td>6.42</td>
<td>&lt;.0001***</td>
</tr>
<tr>
<td>Cosine(Hour Played)</td>
<td>-19.07</td>
<td>11.37</td>
<td>-1.67</td>
<td>0.095</td>
</tr>
<tr>
<td>Device(Touch):log(Trial)</td>
<td>-33.55</td>
<td>9.97</td>
<td>-3.37</td>
<td>0.0008***</td>
</tr>
<tr>
<td>Device(Touch):Age</td>
<td>-10.15</td>
<td>23.58</td>
<td>-0.43</td>
<td>0.67</td>
</tr>
<tr>
<td>Log(Trial):Age</td>
<td>-10.09</td>
<td>4.33</td>
<td>-2.33</td>
<td>0.02*</td>
</tr>
<tr>
<td>Device(Touch):Sex(Male)</td>
<td>-13.84</td>
<td>46.56</td>
<td>-0.3</td>
<td>0.77</td>
</tr>
<tr>
<td>Log(Trial):Sex(Male)</td>
<td>40.97</td>
<td>8.86</td>
<td>4.63</td>
<td>&lt;.0001***</td>
</tr>
<tr>
<td>Device(Touch):log(Trial):Age</td>
<td>0.99</td>
<td>7.15</td>
<td>0.14</td>
<td>0.89</td>
</tr>
<tr>
<td>Device(Touch):log(Trial):Sex(Male)</td>
<td>15.81</td>
<td>14.12</td>
<td>1.12</td>
<td>0.26</td>
</tr>
</tbody>
</table>

3.4 Device type does impact overall skill acquisition

However, the mixed effects model did demonstrate a significant effect of device type on skill acquisition ($\beta_{\text{Device(Touch)}:\text{Trial}} = -33.55$, $t(271) = -3.37$, $p = .0008$). This indicated that participants who performed the task on a desktop improved their performance on the task at a faster rate than those who performed the task on a mobile device. There was also a main effect of device ($\beta_{\text{Device(Touch)}} = -121.81$, $t(271) = -3.63$, $p = .0003$), which also indicated that a participant's initial performance was slower when performing the task on a mobile device than on a desktop computer. We emphasize, however, that these differences in skill acquisition and initial performance are due solely to the device and are not a by-product of age or sex effects (see above results), making it easy to control for in future studies. Full visualization of game performance stratified by device and sex can be viewed on Figure 3.
Figure 3. Average time in target across trial number stratified by participant sex (male = dashed line; female = solid line) and device used (computer = orange; mobile = green).

4. Discussion

The purpose of this study was to investigate the effect of internet device type (computer or mobile) on motor skill acquisition on Super G, a motor-cognitive game designed for use in online cohorts for Alzheimer’s disease research [19]. Age and biological sex are the two highest risk factors for Alzheimer’s disease [20], [21]; however, with the growing trend of digital assessments of cognitive abilities [8], [9], [25], it is crucial to understand how these factors may affect the performance of such digital “games” based on the type of device used, since individuals may be using either a computer (desktop or laptop) or a mobile device (smartphone or tablet) to complete them. The results of this study are encouraging, showing that device type had no effect on Super G based on age or sex, suggesting that Super G is equally sensitive to sex and age regardless of which device type is used for playing.

However, there was a significant main effect of device type on Super G, which is consistent with previous research on remote assessment of other cognitive-motor tests, such as a simple visual reaction time task [26], where better performance is seen when played on a computer. This is an important finding as remote, unsupervised assessments become more prevalent, and likely on mobile devices [10], [14], [27]. While it is critical to understand how the context of the assessment may impact performance (and affect the interpretation of results), it is also easy to control for in post-processing once the degree of offset based on device type is known. For example, in this study the mean difference between mobile and computer performance was 121 ms. Thus, to estimate a person’s mobile performance as if they played on a computer, we would simply add 121 ms to their mobile performance. This type of process is similar within the context of task performance measured across different computer-generated simulations of a three-dimensional environment (virtual reality versus augmented reality) [28].

The results of this study indicate that age impacted performance on Super G, with older adults acquiring the skill at a slower rate than younger adults. This finding differs from previous research suggesting that performance declines with age, but learning capabilities remain relatively intact, i.e., the relative improvement across age may vary, but the rate at which they reach that level is consistent [29]. However, it should be noted that this study did not account
for potential differences in task familiarity. Younger participants may have had more experience with this type of video game task, which could have impacted their performance. Solum and colleagues [30] argued that poorer performance on computer games among older adults could be attributed to previous experience rather than age-related changes. Additionally, recent evidence has suggested that reduced visuospatial memory, which can decline with age, may also impact the rate of motor skill learning among older adults (Wang et al., 2022). Thus, future studies should explore the influence of previous experience and visuospatial memory on the rate of skill acquisition, independent of age.

The results of this study also showed that sex influenced performance in the game, with males performing better than females. This effect is not unique to Super G, but is instead consistent with prior literature. For example, sex differences on a table tennis video game demonstrated that males outperformed females even when matching across previous video game experience [31]. This could be due to biological factors, as previous research has found similar results in visuospatial tasks [32], [33] whereby males tend to perform better on visuospatial assessments than females. The estimated sex effect of this study is equivalent or even larger than that estimated by a previous meta-analysis [34]. We note, however, that in general differences in video game performance between males and females may be due to differences in strategy rather than innate spatial ability [35]. The observed sex effect could also be influenced by the fact that males play video games more often and have more experience [36]. Given that this study was hosted on MTurk and advertised as a video game study, it is possible that this study (unintentionally) selectively recruited women who regularly play video games participated in this study. It is unclear whether the observed sex effect in this study is due to biological, cultural, or a combination of both factors. We also note that future studies should also survey participants on their video game experience.

4.1 Limitations

The study lacked information on the specific device models used by the participants, which could have influenced their gameplay experience. Furthermore, we did not collect data on device screen size or participant internet connectivity, which may also impact performance. It will be important for future research to collect a broad range of device data to limit potential device-specific performance variability which may mask the important person-specific performance variability. Although the overall impacts of technology type may be minimal, i.e. type of browser the game is played, as previous research has shown that modern web platforms provide reasonable accuracy and precision [37]. Additionally, the screen refresh rate was found to be related to individual performance, but it did not appear to impact the effect of device type (See Table 2). Future studies that measure motor performance across different device types should attempt to reduce any between-device performance differences, or at minimum control for device type within their statistical analyses. We also caution against the generalizability of these findings to the overall population at this time, as the study was only conducted on participants recruited through MTurk. Previous research has argued that MTurk “workers” do not necessarily represent the general population [38]. To overcome these limitations, future studies should aim to gather more detailed information on the participants’ devices and use a wider range of recruitment methods to obtain a more representative sample [25].

5. Conclusions

Overall, this study demonstrates that device type did not significantly impact performance of an online motor-cognitive game due to a participants age or sex. There was an overall effect of device type, which should be considered moving forward with future studies of remote, unsupervised motor assessments. Significant effects of sex and age (regardless of device type)
may indicate possible biological, cultural, and aging factors that impact acquisition of a complex motor skill.

**Acknowledgments**

We would like to thank the participants of this study for their individual contribution to this work and the National Institute of Aging [F32 AG071110] for funding and support of this study.

**Conflicts of interest**

The authors have no conflicts of interest to declare.

**References**


