VR FOR THE ELDERLY: QUANTITATIVE AND QUALITATIVE
DIFFERENCES IN PERFORMANCE WITH A DRIVING SIMULATOR

Authors

1. Lili Liu, Ph.D., Associate Professor, Department of Occupational Therapy, Faculty of Rehabilitation Medicine, University of Alberta

2. Benjamin Watson, Ph.D., Assistant Professor, Department of Computing Science, Faculty of Science, University of Alberta

3. Masako Miyazaki, M.Sc., Associate Professor, Director, Telehealth Technology Research Institute, University of Alberta

Source of work: University of Alberta

Correspondence should be addressed to:

Lili Liu, Ph.D., Associate Professor
Department of Occupational Therapy, Faculty of Rehabilitation Medicine
University of Alberta, 3-14 Corbett Hall
Edmonton, Alberta, Canada. T6G 2G4
Phone: (403) 492-5108 Fax: (403) 492-1626
e-mail: lili.liu@ualberta.ca
ABSTRACT

This article presents an analysis of age differences in virtual reality (VR) use. 148 users, grouped by age and balanced by gender, answered questionnaires about their driving, educational, and medical histories, as well as their experience with computers. Participants then completed a driving assessment in a low cost, PC based virtual environment known as DriVR. Dependent measures included both automatic and observer recorded quantitative measures of driving performance, as well as participant reported qualitative measures of difficulty, comfort, visual clarity, delay and simulator sickness. There were few age-related differences in quantitative measures, while qualitative differences between middle aged and elderly participants were seldom found. These results indicated that use of VR with the elderly is quite feasible. In a contradiction of previously reported results, symptoms of simulator sickness increased with age.
INTRODUCTION

As reducing cost increases the application of VR technology in medicine, increasing numbers of clinicians and researchers are examining the use of VR for the elderly. In [1], we reported on the effectiveness of DriVR, a PC-based VR driving simulator, for use in driving assessments of normal and head-injured populations. In this report, we expand on our earlier analyses with detailed consideration of qualitative measures, in an effort to evaluate the more general question of the feasibility of VR use with the elderly.

Computer graphics, VR, and simulation technology have seen only very limited application with the elderly population. As medical success and the baby boom increase the size of the elderly population, this will change. Researchers realize this, and have just begun to examine improvements in the human-computer interface for the elderly [2, 3, 4]. High end simulation technology has traditionally been applied to the training of young pilots and operators, but this too is now changing [5,6]. VR applications for the elderly are few and in their earliest stages [7].

One of the primary concerns in any application of VR technology is simulator sickness. This sickness is thought to result from differences between the perceptual display offered by the VR, and the input expected by the human perceptual system. When the mismatch between display and expected input is particularly strong, symptoms of sickness result [8]. Susceptibility to simulator sickness varies widely and is thought to vary with a number of factors, including gender, experience with the interface, field of view, feedback delay and illness [8, 9, 10, 11]. Age is also thought to be related to simulator sickness, though research reports on this relationship are few [9, 10, 11].
Results indicate that susceptibility of infants to simulator sickness is quite low. From the age of two until puberty, susceptibility increases quickly. Thereafter susceptibility decreases rapidly until 21 years of age. Likelihood of simulator sickness then drops gradually, disappearing almost completely by age 50.

The remainder of this report reviews the components and function of the DriVR simulator and the methodology of our evaluation of it. We then analyze age differences in quantitative and qualitative measures of performance in detail, and conclude with the implications of this analysis for VR use by the elderly.

**DESCRIPTION OF THE DriVR SIMULATOR**

This section is an overview of the DriVR system and its function in hardware and software. See [1] for a more detailed review of the system.

**Hardware**

The DriVR simulator used in this study was a low cost PC-based system. Experimental participants guided the simulator using consumer-level steering wheel and foot pedal input devices. Visual display was provided by the inexpensive Virtual I/O i-glasses, which provide a 30 degree horizontal field of view. The display was used in a biocular mode, with the same image sent to each eye. The i-glasses included three degree of freedom head tracker (yaw, pitch and roll). For audio display, DriVR used a Sound blaster card, with output sent to earphones in the head-mounted display.

**Content & Performance**
DriVR provided one practice and 10 testing scenarios, which appear in a continuous sequence as the participant drives through a small town roughly 1.4 km square. Traffic signs and marks were modeled with textures on simple polygonal objects. Only the participants’ vehicle generated sound: a skidding sound, a collision sound, and an engine sound that varied with engine RPM, to give auditory feedback for speed. Mean frame rate was 14 Hz, and depending on the visual complexity of the current environment, varied between 12 and 18 Hz. Lag in head tracking was not measured but was quite apparent when participants’ turned their heads quickly.

Dependent Measures

The DriVR system provided several measures to aid in assessment of driving. Continuous measures of performance included mean speed, mean distance from the center lane and leading car, the number of incorrect stops, and the number of time the center lane and shoulder were crossed. Binary measures of success in several scenarios included encounters with crossing traffic, a car backing out of a driveway, a repair crew flagman, traffic cones, and a turn into an empty driveway.

Is the DriVR system typical?

Most VR systems in the past years have been based on fairly expensive graphics workstations. Accompanying head mounted displays and tracker systems were also quite dear. If these symptoms had remained the only option, recent interest in the medical community in VR would likely not be so widespread. The recent development of PC-based systems makes VR technology applicable at much lower cost and complexity.
DriVR is exactly this sort of system. Thus, while the DriVR system is not necessarily at this point “typical”, an evaluation of the usefulness of VR systems like DriVR for people of various ages is of particular importance.

**EVALUATION METHODOLOGY**

Each participant was given a pre-test questionnaire. The questionnaire addressed demographics (age, gender, years of education), driving history (car type, driving frequency, license type, driving difficulties), health status (corrective lenses, cognitive change, pre-existing health conditions, current prescriptions) and technology use (computer use and game play frequency).

Participants then began the driving assessment. The assessment began with ten minutes of practice (in a virtual environment slightly different from the testing environment). Participants then completed one driving test. After a five minute rest, the same test was performed again. This second test allowed measuring for the effects of practice. This entire procedure was completed by participants within an hour.

During testing, the evaluator observed the participant closely, recording any failures to check to the left or right at intersections. Participants were monitored for signs of simulator sickness. If any participant experienced severe signs of sickness or physical discomfort, testing was halted, the participant’s data was discarded, and a replacement participant recruited.

After completion of the entire testing procedure, each participant was given a questionnaire to obtain self-reported qualitative results. These results included:
**Difficulty Rating:** participants rated the general difficulty of the test using a 10 centimetre visual analogue scale, zero and ten, zero representing least difficulty.

**Judging distance:** whether or not participants asked about the simulator software included an unprompted complaint about difficulty judging distances.

**Reading signs:** whether or not participants asked about the simulator software included an unprompted complaint about difficulty reading traffic signs.

**Head Mounted Display (HMD) fit:** whether or not participants asked about the HMD included an unprompted complaint about fit.

**HMD With Glasses:** whether or not participants asked about the HMD included an unprompted complaint about fit or discomfort with glasses.

**Head delay:** whether or not participants asked about the HMD included an unprompted complaint about response to head motion.

**Screen resolution:** whether or not participants asked about the HMD included an unprompted complaint about the resolution of the HMD screen.

**See to side:** whether or not participants asked about the HMD included an unprompted complaint about difficulty seeing to the side.

**Foot pedal spacing:** whether or not participants asked about the foot pedals (brake, gas) included an unprompted complaint about the separation between the pedals.


Pedal delay: whether or not participants asked about the foot pedals (brake, gas) included an unprompted complaint about the simulator’s response to the pedals.

Steering wheel size: whether or not participants asked about the steering wheel included an unprompted complaint about inappropriate size.

Wheel delay: whether or not participants asked about the steering wheel included an unprompted complaint about the simulator’s response to the steering wheel.

Simulator sickness: participants were explicitly asked if they experienced any of the following symptoms associated with simulator sickness: discomfort, dizziness, nausea, disorientation, headache, vomiting, or any other symptoms. Severity of reported symptoms were rated by the experimenter as either mild or severe.

AGE DIFFERENCES IN VR USAGE

Objective

The objective of this study was to find how a typical driving population would perform on the DriVR simulator. In this analysis, we focus on differences in quantitative and qualitative effects across age groups, and the relationships between simulator sickness and other participant characterizations.
Participants

All our participants lived in the Edmonton metropolitan area, and were recruited through word of mouth and newspaper advertisements. Participants were grouped into the following eight age categories: less than 16, 16-25, 26-35, 36-45, 46-55, 56-65, 66-75, 76 and older. Twenty participants completed testing in all groups except the oldest, in which eight participants completed testing. Few seniors over the age of 75 were able to travel independently to the test site in Edmonton’s severe winter weather.

Of a total of 162 participants, 148 completed testing. 14 participants were unable to complete testing due to simulator sickness or physical discomfort. The ages of these participants were evenly distributed across the eight age categories. These participants have not been included in the following analyses and were evenly distributed across age groups (see [1]).

For the purposes of this analysis, participants were organized into three groups: less than 36, 36-55, and 56 and older. We attempted to have equal numbers of men and women in each age group. The number of men and women in each of the three groups is shown in Figure 1.

Table 1 shows the driving and educational histories of the participants by age group. All of these participant characterisations had a significant relationship to age group (Pearson Chi Square analyses, driving difficulties p = .01, all others p < .001). The youngest group did not have drivers' licences, therefore, they drove much less than the older groups, while the middle aged group had a higher level of education than the other two groups.
Figure 2 compares frequency of computer use and computer game play to age group. Frequency of computer use declined significantly as age increased (Pearson Chi Square (14) = 70.961, p < .001), as did frequency of computer game play (Pearson Chi Square (16) = 41.225, p = .001). Computer use was still frequent among the middle aged group, while game play was quite infrequent.

Figure 3 compares age group to use of corrective lenses (e.g. glasses, contacts), reports of cognitive change, health conditions, and current use of medication. All the participant characterisations had significant relationships to age group. Just over half of participants in the youngest age group used corrective lenses, but almost all members of the oldest age group used them (Pearson Chi Square (8) = 32.512, p < .001). Almost none of the youngest group reported cognitive changes, while fully half of the oldest group did (Pearson Chi Square (8) = 41.969, p < .001). Reported health conditions (Pearson Chi Square (8) = 23.374, p < .005) and use of medication (Pearson Chi Square (2) = 7.507, p < .05) also increased with age.

Discussion. These characterisations of the study participants with respect to age reflect well known trends that may have implications for VR use. For driving simulations, one can expect that younger users will be less experienced drivers. For more general VR applications, older users will likely be less familiar with computers in general and computer games in particular. At the same time, unlike younger participants, they are probably dealing with many medical and health concerns. Below, we examine the relationship of age and some of these participant characterisations to quantitative and qualitative measures of performance.
Quantitative dependent measures

For all pass/fail measures in the DriVR simulator, performance had no significant relationship to age group. On the other hand, many continuous measures were significantly related to age group. Most of these relationships could be attributed to the tendency of older participants to drive at slower speeds. A typical example of this is the significant relationship between test completion time and age group (one-way ANOVA, F(2) = 31.05, p < .001), diagrammed in Figure 4.

Age and the left/right check measures (recall that it was not an automatic measure in the DriVR simulator) were also significantly related. The oldest group of participants missed significantly more checks than did the younger participants (for left, Pearson Chi Square (12) = 38.48, p < .001, for right, Pearson Chi Square (12) = 33.045, p = .001). These relationships are diagrammed in Figure 5.

Ratings of perceived difficulty for each age group are shown in Figure 6. Ratings are fairly high, and increase significantly with age (one-way ANOVA, F(2) = 4.68, p < .05).

Discussion. The lack of any differences in pass/fail measures across age bodes well for VR applications for the elderly, particularly for simulations of everyday activity with relatively familiar interfaces (e.g. driving). Differences in testing time and driving speed might be attributed to the usual caution that comes with age, as well as the relative unfamiliarity of older users with computer games and interfaces. Unfamiliarity with the interface might also account for the oldest group’s higher left/right check failure rate; many users unfamiliar with head-mounted displays do not realise that moving their head will change the current view. The reduced useful fields of view (UFOVs) [] exhibited by
the elderly in real-world driving may also play a role. Difficulty ratings for the oldest group were no higher than those for the middle aged group.

**Qualitative dependent measures**

Complaints about the fit and comfort and comfort of the VR interface are displayed in Figure 7. Complaints about using the HMD with glasses increased significantly with age group (Pearson Chi Square (2) = 14.53, p = .001). Complaints about other fit issues were not significantly related to age, although the relationship of pedal spacing and age group approached significance (Pearson Chi Square (2) = 4.41, p = .11).

Despite the relatively poor quality of this PC-based VR interface, well under half of the participants complained about visual quality. These complaints, graphed in Figure 8, did not show any significant relationship to age group.

Complaints about delay in its various forms were quite high (Figure 9). (Recall that these complaints might indicate that response might be too slow or too fast). Over 80 percent of participants in all age groups complained about response to the steering wheel. Reports of this type of delay were not significantly related to age group. Over 50 percent of participants in all age groups complained about pedal response. The relationship between age group and pedal response complaints approached significance (Pearson Chi Square (2) = 4.34, p = .11), with the middle aged group complaining most often. Complaints about head delay occurred less frequently, but the relationship between complaints about head delay and age group was marginally significant (Pearson
Chi Square (2) = 5.28, p < .1), with middle aged participants again complaining most frequently, and the oldest group complaining least often.

Reported symptoms of simulator sickness (Figure 10) increased significantly with age (Pearson Chi Square (16) = 27.35, p < .05). Well under half of the participants in the youngest age group reported symptoms, while more than half of the participants reported symptoms in the other groups. The number and severity of these symptoms also seemed to increase with age (Figure 11).

We also analyzed the relationship between reports of simulator sickness and other characteristics of participant population and performance. We found a marginally significant relationship between reported simulator sickness and gender (Pearson Chi Square (2) = 5.11, p < .1), with the number of women reporting severe symptoms being much higher than the number of men reporting such symptoms (see Figure 12). The relationships between reported simulator sickness and frequency of computer use (Figure 13), complaints about head delay, and the ability to see to the side (Figure 14) were not significant. None of these analyses included those participants unable to complete testing due to simulator sickness.

We also examined the relationship between simulator sickness and the following: medication usage, vomiting, and perceived difficulty as rated on the visual analogue scale. Of the ten subjects that were on at least one medication, seven reported having symptoms of simulator sickness (nausea) - five mild and two severe. One subject vomited and, as expected, this subject reported experiencing severe symptoms of simulator sickness. The mean ratings for perceived difficulty among subjects who reported no symptoms, mild symptoms and severe symptoms were as follows: 6.12
(SD=1.97), 6.95 (SD=1.89), and 7.07 (SD=1.79). These differences were statistically significant when compared using a one-way ANOVA (F=3.74, df=2, p=0.026).

Discussion. These results again indicate that using VR with the elderly is quite feasible. Complaints about fit were relatively few. A significant relationship between age and complaints about glasses and the HMD indicate that care should be taken when selecting HMDs for use by the elderly. Interestingly, complaints about visual quality of the VR interface were relatively few, when compared to the number of complaints about delay – apparently, good responsiveness is far more important to users than visual quality. A marginally significant relationship indicated that the elderly are in fact least affected by head delay. Simulator sickness increased significantly with age. However, symptoms reported by the older group were no greater than those reported by the middle aged group. Ratings of perceived difficulty with the DriVR increased with severity of symptoms of nausea.

GENERAL DISCUSSION

As virtual reality becomes cheaper and simpler to use, its use in medical applications will become more common. Since a large portion of the medical patient population is elderly, studies of the use of VR with elderly populations are crucial. Despite this fact, research on the use of VR with the elderly are quite sparse. This study fills something of this gap, reporting on age differences in performance of 148 participants in a low cost VR simulator.
Our results indicate that VR is an appropriate interface for the elderly. Quantitative tests of success or failure did not differ by age group, despite differences among age groups in health condition and familiarity with technology. Differences in continuous measures such as speed seemed to mirror real world differences. Qualitative measures showed few differences between the oldest and middle aged groups. In particular, symptoms of simulator sickness reported by the middle aged and oldest groups were similar. Among the few exceptions, a relatively high number of complaints about fit with glasses indicated that HMDs for the elderly should be selected with care.

Surprisingly, the elderly had fewer complaints about head delay than other groups. This might be accounted for by their slower driving speed – slowing the pace of change is one sure way to compensate for delay in feedback.

The limited amount of previous research on the relationship of simulator sickness to age indicates that susceptibility to simulator sickness increases sharply until puberty, after which it declines sharply until age 21, and is almost completely gone by age 50. Our results are in complete contradiction to this: reported symptoms increased significantly with age. In an effort to explain this, we examined factors previous research has related to simulator sickness susceptibility. Our results confirm the previously reported effect for gender, but since age groups were balanced by gender, an interaction of gender with age cannot provide an explanation. Head delay, the ability to see to the side (field of view), and frequency of computer use (familiarity with the interface, locus of control) all had insignificant relationships to reported simulator sickness, and also could not explain the contradiction.
Explanations for these differing experimental relationships between age and simulator sickness might include limitations of and differences between this and previous studies. Most other experiments [10,11] are already several years old, and generally used higher-end, higher fidelity systems, most without HMDs. Also, this experiment used participant-reported measures and a large scale, normative design that was not meant to address only questions about simulator sickness. Ultimately, resolution of this contradiction will have to await further experimentation on this poorly addressed question.

ACKNOWLEDGEMENTS

This study was funded by a grant from Imago Systems, Inc., Vancouver, British Columbia.

REFERENCES


Table 1: Driving and educational history of experimental participants, by age group.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>13-35</th>
<th>36-55</th>
<th>56+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Own at least one car</td>
<td>71.19%</td>
<td>97.56%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Drive once/wk or more</td>
<td>47.46%</td>
<td>95.12%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Possess valid license</td>
<td>67.80%</td>
<td>100.00%</td>
<td>100.00%</td>
</tr>
<tr>
<td>Have difficulty driving</td>
<td>6.78%</td>
<td>29.27%</td>
<td>29.79%</td>
</tr>
<tr>
<td>Univ education or more</td>
<td>42.11%</td>
<td>70.00%</td>
<td>46.81%</td>
</tr>
</tbody>
</table>
Figure 1: Number of experimental participants in each age group, by gender.

Figure 2: Mean weekly frequency of computer use and game play reported by participants, by age group.

Figure 3: Percentage of each age group reporting use of corrective lenses for vision, changes in cognitive function, current health conditions, and current use of medication.

Figure 4: Mean completion time of the DriVR simulation assessment, by age group.

Figure 5: Mean number of missed left and right turning checks, by age group.

Figure 6: Mean reported difficulty rating, by age group.

Figure 7: Percentage of participants complaining about use of the HMD with glasses, about head-mounted display fit, the size of the steering wheel, and foot pedal spacing, by age group.

Figure 8: Percentage of participants complaining about display resolution, and reporting difficulty judging distance, reading signs, and seeing to the side, by age group.

Figure 9: Percentage of participants complaining about head, pedal, and steering wheel delay, by age group.

Figure 10: Percentage of participants reporting symptoms of simulator sickness of various types, by age group.

Figure 11: Percentage of participants reporting symptoms of simulator sickness of varying severity, by age group.

Figure 12: Percentage of male and female participants, by reported symptoms of simulator sickness.
Figure 13: Mean frequency of computer use and game play, by reported symptoms of simulator sickness.

Figure 14: Percentage of participants complaining about head delay and seeing to side, by reported symptoms of simulator sickness.
Age Groups (Yrs)

Mean Frequency (Days/Wk)

- comp use
- game play

13-35
36-55
56+
Age Groups (Yrs)

Participants Complaining (Pctg)

- screen res
- judge dist
- read signs
- see to side

13-35
36-55
56+
Gender Composition (Pctg)

Sim Sickness Symptoms

male
female

none  mild  severe
Sim Sickness Symptoms

Freq Computer Use (Days/Wk)

none    mild    severe

0   1   2   3   4   5   6   7
Cyberpsychology & Behavior, 2, 5

\[ \text{Sim Sickness Symptoms} \]

![Bar chart showing participants complaining (Pctg) for different levels of sim sickness symptoms: none, mild, severe. The chart compares head delay (black bars) and see to side (gray bars).]