



Simulator sickness during driving simulation studies

Johnell O. Brooks^{a,*}, Richard R. Goodenough^a, Matthew C. Crisler^a, Nathan D. Klein^a, Rebecca L. Alley^a, Beatrice L. Koon^a, William C. Logan Jr.^c, Jennifer H. Ogle^b, Richard A. Tyrrell^a, Rebekkah F. Wills^a

^a Clemson University, Department of Psychology, 418 Brackett Hall, Clemson, SC 29634, United States

^b Clemson University, Department of Civil Engineering, 116 Lowry Hall, Clemson, SC 29634, United States

^c Greenville Hospital System University Medical Center, Division of Geriatrics, 255 Enterprise Boulevard, Suite 101, Greenville, SC 29615, United States

ARTICLE INFO

Article history:

Received 13 August 2008

Received in revised form 1 December 2008

Accepted 21 April 2009

Keywords:

Simulator sickness

Motion sickness

Driving

Simulation

Simulator

ABSTRACT

While driving simulators are a valuable tool for assessing multiple dimensions of driving performance under relatively safe conditions, researchers and practitioners must be prepared for participants that suffer from simulator sickness. This paper describes multiple theories of motion sickness and presents a method for assessing and reacting to simulator sickness symptoms. Results showed that this method identified individuals who were unable to complete a driving simulator study due to simulator sickness with greater than 90% accuracy and that older participants had a greater likelihood of simulator sickness than younger participants. Possible explanations for increased symptoms experienced by older participants are discussed as well as implications for research ethics and simulator sickness prevention.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Simulation is an invaluable research tool. Not only can it produce scenarios that are logistically or monetarily impossible, but it also eliminates a great deal of the risk associated with performing tasks in the real world. For this reason, although nearly any task can be simulated, simulation is most often used for tasks involving some degree of danger in order to provide an avenue for training, research, and even entertainment. It should come as no surprise then that some of the most commonly simulated tasks include flight, medical procedures, and the focus of the current study, driving.

Although simulation can eliminate the crash risks associated with on-road research, the use of simulation introduces another risk, a syndrome known as simulator sickness (SS). This malady, similar to motion sickness (MS), can potentially confound data (Lerman et al., 1993; Cobb et al., 1999), limit the effectiveness of training (Hettinger et al., 1990), and influence participant dropout rates (Cobb et al., 1999). This article addresses the various theories of MS and SS as well as some common measurement scales. Next, this article will present a method used in the Clemson University driving simulator laboratory to protect participants from simulator sickness. Finally, this article will consider practical concerns for practitioners and researchers dealing with simulator sickness and ways in which these concerns may be addressed.

1.1. The relationship between motion sickness and simulator sickness

Like MS, SS has been described as a syndrome because of the breadth of its symptoms, including headache, sweating, dry mouth, drowsiness, disorientation, vertigo, nausea, dizziness, and vomiting (Ebenholtz, 1992; Kennedy et al., 1993; Cobb et al., 1999). Cobb et al. (1999) have also documented a negative effect on psychomotor control, believed to be the product of SS. Moreover, user characteristics such as age, experience, gender, illness, mental rotation ability, and postural instability play key roles in determining whether a participant will become sick.

Older adults tend to be more susceptible to SS than younger participants (Roemaker et al., 2003). Additionally, SS may vary by exposure time; Cobb et al. (1999) have suggested that SS symptoms steadily increase for up to one hour during exposure to a virtual environment before returning to nominal levels 15 min later. During this adaptation period, however, some subjects may become too ill to continue and thus never reach the 1-h mark. Finally changes in scene content may affect the likelihood and severity of SS (Jones et al., 2004).

While some researchers view SS as a type of MS which occurs in a simulated environment, there are several reasons to treat MS and SS as related but separate maladies. To begin with, MS appears to occur in a larger portion of the population and tends to be more severe than SS. Additionally, a key indicator of MS, drowsiness, does not necessarily indicate SS (Kennedy et al., 1993). Furthermore, eye movement disturbances are more common in SS.

* Corresponding author. Tel.: +1 864 656 2831.

E-mail address: jobrook@clemson.edu (J.O. Brooks).

1.2. Theories of motion sickness and simulator sickness

Over the years, researchers have developed numerous theories explaining how MS and SS occur. The three most widely accepted by the MS and SS research community are the sensory conflict, postural instability, and eye movement theories. A fourth theory, the evolutionary theory, explains why, rather than how, MS and SS occur.

1.2.1. Sensory conflict theory

Reason and Brand's (1975) sensory conflict theory is probably the most widely accepted theory of MS and SS. The theory proposes that a conflict between or within sensory systems causes MS symptoms to arise. Specifically, conflicts between the motion one sees and the actual motion one is experiencing as well as conflicts between the structures within the vestibular system which detect and perceive direction and acceleration of motion are the two main contributors to MS and SS (Reason and Brand, 1975; Regan, 1994; as cited in Cobb et al., 1999).

In 1978, Reason proposed a Neural Mismatch Model suggesting that, for sickness to occur, sensory information must also be in conflict with one's own past experiences of a motion environment. Based on this model, sickness is most likely when sensory information is repeatedly contradictory, greatly disparate, or does not match one's expectations. More recently, Bles et al. (1998) have suggested that the visual–vestibular conflict is necessary and sufficient to produce motion sickness.

1.2.2. Postural instability theory

Riccio and Stoffregen (1991) opposed the sensory conflict theory by noting that congruent information from sensory systems is unusual even in normal, everyday tasks. Instead, they point out that maintaining postural stability is a natural inclination in most animals. According to this theory, MS occurs when one is placed in a novel environment in which effective ways to maintain balance have not been learned (Duh et al., 2004). For example, travelers at sea must learn ways to adjust to a ship's motion, often referred to as getting one's "sea legs." Once they return to land, their sea legs come with them, sometimes causing them to sway when standing or walking until they adapt to being back on land.

1.2.3. Eye movement theory

According to the eye movement theory of MS, certain stimuli can cause eye movements which create such tension in the eye muscles that they stimulate the vagus nerve resulting in MS (Ebenholtz, 1992). Ebenholtz (2001) has proposed that two specific eye movements, optokinetic nystagmus and vestibular ocular response, lead to MS and SS. In optokinetic nystagmus the eye pursues a target object from one end of a visual scene to the other. When the eye can pursue the object no further, it snaps back to the far side of the visual field where it begins to pursue again. Similarly, the vestibular ocular reflex is responsible for keeping a target object on the fovea (i.e. the center of the retina where one's vision is sharpest) when the head is turning. Thus, if one rotates one's head to the right 3° while fixating an object straight ahead, the vestibular ocular reflex causes the eye to rotate to the left 3°. Errors in these eye movements can result in headache, eye strain, and difficulty concentrating.

1.2.4. Evolutionary theory

Treisman's (1977) evolutionary theory of MS differs from the three aforementioned theories in that it attempts to explain why MS and SS occur rather than how they occur physiologically. Specifically, Treisman suggests that the human species has not had sufficient time to adapt to the relatively new modes of transportation we use today and that the body responds to conflicts in sensory information as if it had ingested poison, the effective reaction being

vomiting, a common MS symptom (Money and Myles, 1974; Money and Cheung, 1983).

1.3. Methods of measurement

Two common surveys measuring symptoms of MS and SS are the Motion Sickness Assessment Questionnaire and the Simulator Sickness Questionnaire (SSQ). Although there are other measures of MS and SS, such as heart rate (Cobb et al., 1999), they are often recorded as secondary data.

1.3.1. Simulator Sickness Questionnaire (SSQ)

The SSQ (Kennedy et al., 1993) is the most frequently used measure of SS. It was developed to replace the Pensacola Motion Sickness Questionnaire (Kellogg et al., 1965) as a measure of the MS-like symptoms sometimes experienced during or after simulator use. The developers of the SSQ felt that a separate measure for SS was needed because of the slight difference in symptoms as well as their lower incidence and severity.

The SSQ contains 16 items rated by participants as "none", "slight", "moderate", or "severe". These items form three subscales, (1) nausea, (2) oculomotor disturbances (such as headache, eyestrain, and blurred vision), and (3) disorientation, which are combined by a series of mathematical computations to produce an overall SS score.

In 2006, Drexler suggested it may be possible to use the symptom subscales of the SSQ to distinguish various types of simulated environments based on the symptoms participants exhibit over a large number of exposures. It could also be possible to make predictions as to what symptoms a given simulator may produce given its attributes and the symptom sets produced by similar simulators.

1.3.2. Motion Sickness Assessment Questionnaire (MSAQ)

The MSAQ (Gianaros et al., 2001), another common measure of MS, asks subjects to rate the severity of four types of symptoms: (1) gastrointestinal (sick to stomach, queasy, nauseated, may vomit), (2) central (faint-like, lightheaded, disoriented, dizzy, spinning), (3) peripheral (sweaty, clammy/cold sweat, hot/warm), and (4) sopite-related (annoyed/irritated, drowsy, tired/fatigued, uneasy). The measure requires participants to rate the degree to which they are experiencing 16 symptoms from 1 (not at all) to 9 (severe). An overall MS score can then be calculated along with a score for each particular domain.

1.4. Use of SS Questionnaire in the Clemson University driving simulation laboratory

When we first began running participants in our driving simulator laboratory, we followed numerous suggestions for preventing SS, including screening for a history of motion sickness, migraine headaches, and pregnancy; keeping the room at a cool temperature; designing studies to allow participants ample opportunity to slowly adjust to the simulator experience; and encouraging participants to express any discomfort they felt during the study. Despite all of these precautions, we were unable to prevent all incidences of SS in the lab. Thus, we decided that a new screening measurement for SS, both before and during the study, was necessary.

This tool had to be quick and easy to administer as well as reliable at predicting SS. Pilot tests using both the SSQ and MSAQ revealed that participants were frustrated with the standard MSAQ rating scale. Specifically, participants with no symptoms repeatedly told us their score should be a "0" instead of a "1" (the lowest score on the MSAQ) since they were not experiencing symptoms while participants with severe symptoms felt their score should be a "10"

rather than a “9” (the highest score on the MSAQ) because they felt 100% of a given item rather than 90%. Therefore, in the remainder of the pilot tests we adopted a 0–10 scale. We have successfully used this altered scale of the MSAQ for a variety of studies to help identify when participants are feeling uncomfortable in the driving simulator. While some may argue that this procedure cues participants to express feelings of SS, participant well-being overrides this concern.

2. Method – exploratory study

2.1. Participants

Three simulator studies (Brooks et al., 2007; Klein and Brooks, 2008; Martin et al., 2007) were combined including a total of 114 licensed drivers. See Table 1 for a description of participant gender, age, and years of driving experience as well as the number of participants who withdrew from each study due to SS. All participants received extra credit in a psychology course or up to \$10/h. None of the participants had previous experience with the driving simulator.

2.2. Driving simulator

Participants drove a medium-fidelity fixed-base driving simulator (DriveSafety, Inc.) consisting of the body of a Mitsubishi Galant sedan and three screens in front of the vehicle. Participants used the sedan’s original steering wheel and pedals. Speed was presented on the screen, using the car’s speedometer, or not at all depending upon the purpose of each study. The roadway environment consisted exclusively of a flat but unusually curvy two-lane rural road (see Fig. 1). Participants never encountered other vehicular traffic; however, some periodically encountered stationary pedestrians standing on either shoulder of the road.

Table 1

Participant age and driving experience (the Brooks et al. (2007), Martin et al. (2007), and Klein and Brooks (2008) studies form the exploratory dataset, and the data from Hein (2007) was used for confirmation of the models).

		N	Males	Age			Years of driving experience			Stopped study
				Mean	SD	Range	Mean	SD	Range	
Brooks et al. (2007)	Younger	14	6	19.3	1.1	18–21	3.7	1.5	2–7	1
	Middle-age	18	10	43.2	4.8	35–50	27.7	5.9	16–38	3
	Senior	22	17	70.7	4.0	65–79	51.2	11.1	12–60	7
Martin et al. (2007)	Younger	17	7	20.0	2.5	18–29	4.5	2.5	2–13	0
	Senior	19	13	71.9	4.1	66–81	53.3	7.5	40–73	5
Klein and Brooks (2008)	Younger	11	6	20.5	1.4	18–23	5.1	1.4	3–7	0
	Senior	13	5	72.2	4.4	66–80	53.5	5.6	47–64	3
Hein (2007)	Younger	25	16	20.3	1.6	18–24	4.6	1.6	2–8	1
	Middle-age	24	11	44.4	6.4	35–62	27.1	3.8	15–41	2
	Senior	24	12	71.5	3.3	65–75	55.7	6.8	49–61	2



Fig. 1. Depiction of a typical scene from the driver’s perspective in the simulator. All scenarios consisted of curvy two lane rural roads free of traffic.

2.3. Procedure

2.3.1. Driving simulator training

Participants were allowed considerable practice time to become comfortable with and competent at using the simulator prior to data collection. The initial 2-min training session consisted of a straight road. The duration of all subsequent training and experimental trials was 5 min. The second practice session introduced the curvy road type that was used in the remainder of the experiment. During the third session, the participants wore either trial frames or lens clips (designed to hold lenses) for the later experimental manipulations. Due to extreme visual challenges experienced during some of the experimental sessions, the drivers practiced error recovery techniques during the fourth training session. Finally, participants completed a baseline session where data were collected to ensure minimal driving performance. Participants were encouraged to repeat each type of training as many times as necessary until they felt comfortable with the task. All sessions were followed by a mandatory 2-min break.

2.3.2. MSAQ

Prior to and after each session, participants completed the MSAQ using the modified scale of 0–10. Participants were asked to state the number that best described how they felt on a scale of 0–10, where 0 is “not at all” and 10 is “severely”. Therefore, data analyses include data from the preliminary MSAQ in addition to data collected after the five training sessions (straight road, introduction of the curvy road, curvy road with trial lens frames, curvy road with error recovery, and baseline trial).

3. Method – confirmation study

MSAQ data from an additional study (Hein, 2007) are used to confirm models created from our three initial studies (Brooks et al., 2007; Martin et al., 2007; Klein and Brooks, 2008). Major differences

between the three exploratory studies above and the confirmation study include the shortening of the training procedure and the use of longer driving scenarios during the experimental trials (30 min in the confirmation study versus five minutes in the exploratory studies). The training procedure was shortened because trial frames were not used. Therefore, there were fewer training scenarios; however, the pattern was the same (straight road progressing to curvy road with a final check to ensure minimal driving performance level).

3.1. Participants

A total of 73 licensed drivers participated in the confirmation study. Drivers belonged to one of three age groups, 17–25 years, 35–64 years, or over 65 years of age. See Table 1 for a description of the participants' gender, age, and years of driving experience. All participants received extra credit in a psychology course or \$10/h for participating in this study.

3.2. Driving simulator

Participants used the same simulator described above. The experimental driving scenarios consisted of a combination of rural and small town roads. The rural areas included straight sections, curves and hills while the roads in town were primarily straight.

3.3. Procedure

3.3.1. Driving simulator training

A subset of the training sessions described in Section 2.3.1 was utilized. Since the visual scenes were not manipulated, two of the sessions were not used (curvy road with frames and error recovery). Therefore, all participants completed a minimum of three training sessions including the straight road, the curvy road, and the final practice session.

3.3.2. MSAQ

Prior to beginning the study and after each driving session, participants were asked to complete the MSAQ. Therefore, the confirmatory analysis includes MSAQ data from the preliminary questionnaire in addition to data collected after the three training sessions (straight road, introduction of the curvy road, and baseline trial).

4. Results and discussion

4.1. Exploratory study data reduction and analysis

The MSAQ datasets from each of the three studies were combined into a single dataset for analysis. The maximal value reported for each question by each participant was then extracted from the complete dataset (the maximum score from the preliminary MSAQ assessment and five training scenarios). For example, if participant 1 reported "queasy" values of 0, 1, 1, 3, 5, and 3 for the preliminary assessment and five training scenarios, respectively, the maximal value would be 5. Note that this includes only MSAQ data from the baseline and training scenarios.

None of the scenarios with experimental manipulations were included; however, if a participant failed to complete the study due to simulator sickness symptoms during the later experimental phase, that participant was coded as such regardless of whether the participant stopped during the training or the experimental phase. For example, if a participant's maximum score in the training scenarios was a 3, but later on reported a 5 for one of the experimental scenarios, the value reported and analyzed here would remain a 3. Descriptive statistics for participants that completed

and failed to complete the experiments and a correlation matrix between all MSAQ item variables and the sickness outcome variable are reported for this dataset to show relationships between individual assessment items as well as their relationships to the sickness outcome variable.

To highlight the importance of the preliminary MSAQ assessment, the correlation analysis was conducted on both the original dataset and a modified dataset that subtracted the preliminary MSAQ value for each question from the associated maximal value reported on that question (e.g. if a participant reported a 3 for the "tired/fatigued" question in the preliminary assessment, and then reported a 7 after one of the driving scenarios, this dataset would have a value of $7 - 3 = 4$ for that question). All subsequent analyses were conducted on the modified dataset consisting of maximum values with preliminary values subtracted in order to avoid associating initial drowsiness, uneasiness, etc. with SS. Although the use of difference scores can be problematic in many instances (Chronbach and Furby, 1970), in this instance the absolute scores for certain items (e.g. "drowsy") are unlikely to represent MS symptoms given that many participants report some degree of drowsiness prior to engaging in the simulation. Simple difference scores were used due to the fact that they are simple to calculate and intuitive for a simulator researcher to use while actually collecting data. For example, if a participant reports being "tired/fatigued" when beginning an experiment, continued reports of being equally "tired/fatigued", are not likely indicative of SS.

A logistic regression analysis was conducted using the maximal MSAQ item scores adjusted for preliminary responses to predict the probability of a participant dropping out of the experiment. A forward conditional entry was used that required $p < 0.05$ for inclusion in the model and $p > 0.10$ for removal from the model. The resulting model was then tested in our confirmatory study using Hein's data (2007).

Since experimenters and participants use these data to subjectively determine if the study should be stopped, we also describe the pattern of responses from each of the participants who withdrew from the study. This descriptive analysis is important in addition to using statistical techniques in order to highlight the diverse patterns observed in our exploratory studies. This allows the reader to understand the variety of patterns experienced by participants who withdrew from our studies.

4.2. Descriptive statistics

Descriptive statistics for the maximum MSAQ value dataset with the preliminary assessment subtracted are shown in Table 2 for both participants who completed the study and for participants who were unable to complete the study. Table 3 shows descriptive statistics for a subset of the individuals from Table 2, specifically for only the senior participants adjusted for preliminary responses. The only individual item with a significant difference observed between senior and non-senior participants was "uneasy", amongst participants that completed the study. For this item, senior participants reported higher values ($M = 1.26$ and 0.12 for senior and non-senior participants, respectively; $t(39.8) = 2.8$, $p = 0.008$). This suggests older participants may be more likely than younger participants to express uneasiness unrelated to SS.

In general, the mean scores for all items were fairly low for both participants that completed and failed to complete the experiment. The highest mean value for any single item was 2.21 for "uneasy", "lightheaded", and "sweaty" for sick participants. However, the range of these variables was 0–8 for "sweaty" and "lightheaded" and 0–6 for "uneasy". This implies that even amongst the participants that failed to complete the study, there were a number of items for which they reported low responses, which kept the

Table 2
Descriptive statistics for participants that completed and failed to complete the experiment adjusted for preliminary response.

MSAQ question	Participants who completed experiment (N=95)				Participants who dropped (N=19)				Δ M2 – M1
	Minimum	Maximum	Mean (M1)	SD	Minimum	Maximum	Mean (M2)	SD	
Sick to my stomach	0	8	0.25	0.99	0	9	2.16	2.50	1.91
Faint-like	0	6	0.19	0.73	0	4	0.42	1.07	0.23
Annoyed/irritated	0	10	0.31	1.17	0	10	1.89	2.98	1.59
Sweaty	0	7	0.29	1.10	0	8	2.21	2.80	1.92
Queasy	0	5	0.40	0.93	0	5	2.37	1.77	1.97
Lightheaded	0	5	0.51	1.04	0	8	2.21	2.23	1.71
Drowsy	0	4	0.22	0.73	0	2	0.11	0.46	-0.11
Clammy/cold sweat	0	5	0.09	0.57	0	10	1.53	2.80	1.43
Disoriented	0	5	0.26	0.70	0	5	1.16	1.50	0.89
Tired/fatigued	0	4	0.24	0.66	0	3	0.58	1.07	0.34
Nauseated	0	5	0.24	0.70	0	8	2.00	2.43	1.76
Hot/warm	0	7	0.32	1.06	0	9	2.18	2.78	1.87
Dizzy	0	6	0.55	1.02	0	8	2.00	2.16	1.45
Like I am spinning	0	5	0.33	0.82	0	5	1.16	1.74	0.83
As if I might vomit	0	1	0.03	0.15	0	10	1.89	2.77	1.87
Uneasy	0	10	0.59	1.74	0	6	2.21	2.02	1.62

means low across participants. This suggests there is likely not a single item or small set of items that will indicate SS for all participants. The means for participants that failed to complete the study were also lower due to the fact that the data presented here include only the training scenarios. There were three seniors that showed little or no symptoms during the training, but eventually reported higher scores during the experimental scenarios. This is important when using the MSAQ to monitor SS because it implies that an experimenter cannot simply compare individual scores to these mean values to determine if a participant should continue a study.

4.3. Correlational analysis

A matrix of correlations between individual MSAQ item scores and sickness outcome is presented in Table 4 for participants of all ages using the maximal value dataset adjusted for preliminary response. Significant correlations are shown in bold. The only items which failed to reveal significant correlations to sickness outcome were “faint-like”, “drowsy”, and “tired/fatigued”. Two of these items, “drowsy” and “tired/fatigued”, are sopite-related symptoms that Kennedy et al. (1993) suggest are less relevant to SS.

After comparing the correlation matrix presented in Table 4 to a correlation matrix on the unadjusted data, notable differences between these matrices were only observed for the “drowsy” and “tired/fatigued” items. The range of mean differences between the adjusted and non-adjusted data was 0 for five MSAQ items

to 0.58 for the “tired/fatigued” item. The mean responses for the “tired/fatigued” item changing by 0.58 and the “drowsy” measure changing by 0.38 were the most notable changes.

As would be expected, some MSAQ item scores were highly correlated (e.g. “hot/warm” and “sweaty”, $r = .827$; “lightheaded” and “dizzy”, $r = .780$; “sick to my stomach” and “nauseated”, $r = .756$). Note that the item scores making up the factors identified by Gianaros et al. (2001) are generally correlated with each other more than with item scores from other factors, and the correlations to sickness outcome are generally strongest for items within the gastrointestinal factor (items “sick to my stomach”, “queasy”, “nauseated”, and “as if I might vomit”) which also explained most of the variance in the Gianaros et al. (2001) factor analysis. Also note that items “faint-like”, “drowsy”, and “tired/fatigued” were not significantly correlated with sickness outcome even after correcting for initial MSAQ responses.

4.4. Logistic regression analysis

The coefficients of two logistic regression models are presented in Table 5. The first model resulted in a non-significant predictor (“sweaty”) that was subsequently removed in model two; however, the classification of participants and the Nagelkerke R^2 , a measure of effect size for logistic regression models, for model one ($R^2 = 0.595$) were better than for model two ($R^2 = 0.585$). Table 6 shows the two models’ performance classifying participants as completing and failing to complete the experiment (assuming a

Table 3
Descriptive statistics for senior participants that completed and failed to complete the experiment adjusted for preliminary response.

MSAQ question	Participants who completed experiment (N=39)				Participants who dropped (N=15)				Δ M2 – M1
	Minimum	Maximum	Mean (M1)	SD	Minimum	Maximum	Mean (M2)	SD	
Sick to my stomach	0	8	0.49	1.48	0	9	2.20	2.65	1.71
Faint-like	0	6	0.36	1.06	0	4	0.53	1.19	0.17
Annoyed/irritated	0	10	0.56	1.73	0	10	2.20	3.23	1.64
Sweaty	0	7	0.41	1.37	0	8	2.53	2.95	2.12
Queasy	0	5	0.64	1.25	0	5	2.40	1.76	1.76
Lightheaded	0	5	0.67	1.30	0	5	1.93	1.87	1.27
Drowsy	0	4	0.23	0.84	0	2	0.13	0.52	-0.10
Clammy/cold sweat	0	5	0.18	0.85	0	10	1.93	3.03	1.75
Disoriented	0	5	0.31	0.92	0	3	1.13	1.25	0.83
Tired/fatigued	0	4	0.28	0.86	0	3	0.73	1.16	0.45
Nauseated	0	5	0.33	0.90	0	8	2.00	2.56	1.67
Hot/warm	0	7	0.49	1.30	0	9	2.53	2.95	2.05
Dizzy	0	6	0.56	1.25	0	5	1.87	1.68	1.30
Like I am spinning	0	3	0.26	0.68	0	5	1.27	1.83	1.01
As if I might vomit	0	0.5	0.01	0.08	0	10	1.87	2.92	1.85
Uneasy	0	10	1.26	2.52	0	6	2.00	2.00	0.74

Table 4
Correlation matrix for maximal MSAQ item scores dataset with adjustment for preliminary values.

	0	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16
0. Sickness outcome																	
Q1. Sick to my stomach	0.468																
Q2. Faint-like	0.109	0.365															
Q3. Annoyed/irritated	0.349	0.242	0.047														
Q4. Sweaty	0.431	0.405	0.397	0.397													
Q5. Queasy	0.556	0.705	0.312	0.312	0.349												
Q6. Lightheaded	0.441	0.456	0.294	0.415	0.459	0.459											
Q7. Drowsy	-0.062	0.066	-0.052	0.330	0.459	-0.075	0.007										
Q8. Clammy/cold sweat	0.399	0.387	0.052	0.330	0.479	0.479	0.305	0.032									
Q9. Disoriented	0.357	0.416	0.357	0.206	0.411	0.411	0.305	0.048	0.192								
Q10. Tired/fatigued	0.168	-0.042	0.181	0.398	0.164	0.038	0.131	0.444	0.067	0.150							
Q11. Nauseated	0.494	0.756	0.284	0.442	0.441	0.690	0.486	-0.032	0.355	0.261	0.149						
Q12. Hot/warm	0.428	0.332	0.174	0.456	0.827	0.400	0.320	0.004	0.716	0.188	0.243	0.479					
Q13. Dizzy	0.394	0.302	0.349	0.168	0.192	0.370	0.780	0.119	0.287	0.478	0.215	0.303	0.265				
Q14. Like I am spinning	0.292	0.075	0.259	0.257	0.292	0.216	0.356	-0.032	0.312	0.143	0.212	0.147	0.479	0.507			
Q15. As if I might vomit	0.532	0.560	-0.061	0.182	0.382	0.536	0.166	-0.056	0.468	0.065	-0.076	0.616	0.348	0.507	0.114		
Q16. Uneasy	0.323	0.450	0.276	0.604	0.323	0.525	0.510	0.226	0.339	0.380	0.362	0.526	0.370	0.348	0.241	0.253	

Bold: significant correlation ($p < 0.05$).

Table 5

Logistic regression model coefficients predicting probability of failing to complete the experiment.

MSAQ item	B	e ^B	SE	Wald	p
Model 1 ($R^2 = 0.595$)					
Sweaty	0.238	1.269	0.224	1.126	0.289
Queasy	0.504	1.656	0.265	3.631	0.057
Dizzy	0.577	1.780	0.221	6.796	0.009
As if I might vomit	1.676	5.342	0.854	3.850	0.050
Constant	-3.558	0.028	0.595	35.819	0.000
Model 2 ($R^2 = 0.585$)					
Queasy	0.492	1.635	0.263	3.506	0.061
Dizzy	0.602	1.826	0.222	7.379	0.007
As if I might vomit	2.045	7.732	0.790	6.709	0.010
Constant	-3.477	0.031	0.580	35.882	0.000

cutoff probability of 50% where any participant that was predicted to be more than 50% likely to fail to complete the study would be classified as failing to complete) for both the exploratory and confirmatory datasets.

The logistic regression models presented in Table 5 suggest that the “as if I might vomit” item has the largest effect on the probability of failing to complete the experiment. Model one includes one item from the peripheral (“sweaty”), one from the central (“dizzy”), and two items from the gastrointestinal (“queasy” and “as if I might vomit”) dimensions of the MSAQ. These are the three dimensions that explained the most variance in motion sickness as identified by Gianaros et al. (2001). Model two includes two items from the gastrointestinal (“queasy” and “as if I might vomit”) and one item from the central (“dizzy”) dimensions. These were the two dimensions that explain the most variance in the Gianaros et al. (2001) factor analysis. Note that neither model includes any items from the sopite dimension, which is consistent with Kennedy et al. (1993).

Both models performed well in terms of classifying participants as completing and not completing the experiment (better than 90% correct for both the exploratory dataset and the confirmation dataset). Some of the misclassifications are likely due to the fact that some participants did not show any symptoms until after the training. It must be noted that SS does not seem to result in the same symptom set for all participants which is problematic for the use of a one-size-fits-all statistical model. This suggests that although the numerical model presented here may be useful, it is still important to pay attention to scores on the other MSAQ items that are not utilized by the model.

4.5. Individual level analyses on participants who withdrew

Four participants under the age of 50 withdrew from the study (three males and one female) while 15 participants over the age of 65 withdrew (nine males and six female). Table 7 identifies the number of items with a non-zero response after the participants’ final session completed for each of the four domains of the MSAQ regardless of whether it was during the training phase or the experimental phase of the study. Table 7 also indicates whether the participants’ scores increased progressively over multiple driving scenarios or suddenly after a single driving scenario. For example if a participant’s scores for queasy were 1, 3, 5, and 8; the participant would be coded as a progressive increase whereas if a participant’s scores were 0, 0, 0, and 3, the participant would be coded as a sudden increase. The session number after which a participant withdrew from the study is also indicated. Session numbers greater than five occurred after the training.

Interestingly, only two of the 19 participants who withdrew from the study had only gastrointestinal symptoms, while seven of the participants had symptoms in all four dimensions. The data also reveals that the pattern in which participants become sick differs.

Table 6
Classification table for logistic regression models 1 and 2 applied to maximum MSAQ item scores dataset.

Dataset	Observed Outcome	Model 1			Model 2		
		Predicted outcome			Predicted outcome		
		Completed	Dropped	% Correct	Completed	Dropped	% Correct
Exploratory	Completed	94	1	98.9	92	3	96.8
	Dropped	7	12	63.2	7	12	63.2
	Overall			93.0			91.2
Confirmatory	Completed	65	3	95.6	64	4	94.1
	Dropped	1	4	80.0	1	4	80.0
	Overall			94.5			93.1

Table 7
Number of items within each dimension of the MSAQ with a non-zero response for participants that withdrew from the study.

Participant #	Under 50 years				Seniors														
	1	2	3	4	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Gastrointestinal (4 items)	1	3	0	4	0	4	4	4	3	4	3	2	2	4	4	1	4	3	0
Central (5 items)	1	0	3	1	3	4	1	3	0	0	3	4	3	0	4	5	3	0	4
Peripheral (3 items)	2	0	0	0	0	2	0	0	3	0	2	3	0	1	3	3	2	0	1
Sopite (4 items)	1	0	1	2	4	2	0	2	0	0	1	2	0	1	2	3	3	1	1
Increase in score sudden (S) or progressive (P)?	S	P	P	S	P	P	P	S	S	S	P	S	P	P	P	P	P	P	P
Final session number	2	2	3	3	2	2	5	17	2	7	2	2	5	3	3	4	2	9	4

Note: Shaded columns indicate participants that logistic regression model one missed classifying as SS.

While some participants' symptoms increase progressively over multiple driving scenarios, others have symptoms that appear suddenly. The variations in these patterns of data show a collaborative effort is necessary between the participant and the experimenter in order to ensure participant well-being. It must also be noted that in some instances participants' scores increased at the beginning of an experiment but subsequently decreased as they adapted to the simulator.

It has been suggested by Kennedy et al. (2000) that sickness severity increases over periods of 20 min (Regan, 1995), 40 min (Lampton et al., 1994), and 60 min (Guignard and McCauley, 1982). It is generally accepted that that incidence of simulator sickness is related to exposure length (Kennedy et al., 2000).

5. Conclusions

While driving simulators are a useful research tool, simulator sickness can be a common side effect. This paper examines the utility of using the Motion Sickness Assessment Questionnaire (MSAQ) with a modified scale during driving simulation studies to determine if a given participant is experiencing SS such that an experiment should be stopped. In this study, a model was generated and tested based on the MSAQ assessment. The model successfully classified greater than 90% of participants as successfully completing or failing to complete the experiment. Although the data presented here were collected in the context of a driving simulator, it is likely that this process would be applicable to other types of simulation.

An analysis of the individual participants who withdrew from the study due to sickness was valuable to illustrate the degree to which it is important to consider each individual participant's data. While the models show that three to four items predict most sickness, the examination of the individual participant's data demonstrate the degree to which the pattern of symptoms/dimensions differs between individuals who became sick. The use of the entire MSAQ facilitates collaboration between experimenter and participant to ensure participant well-being. Additionally, labs must also consider other issues including: the age of the participants, ethical issues, and research review boards.

5.1. Older participants

These data suggest that older participants have more difficulty with SS than younger participants. One possible explanation is the increased number of balance and dizziness problems experienced with aging. Causes can range from clogged or narrow ear canals to posterior cerebral circulation strokes to simple loss of sense of position (Von Brevern et al., 2007). Despite being common, the medical literature lacks systematic evaluation of balance and driving issues, making it difficult for physicians to make informed statements about the frequency, severity or significance of dizziness in older patients (Ardic et al., 2006).

Anti-cholinergic medications like Dramamine, Meclizine, and Scopolamine have been evaluated as treatments for MS with variable results (Karkos et al., 2007). These evaluation studies have included few older patients. In clinical experience, these medicines frequently cause disturbing central nervous system side effects like sleepiness, memory lapses, urinary retention, and even delirium when used in older patients (Burke, 1995). These common side effects in older adults make the use of these classes of medications as preventive therapy difficult if not dangerous.

5.2. Research ethics and institutional review

From an Institutional Review Board (IRB) perspective, the possibility of simulator sickness raises several issues for consideration. Specifically, SS may be considered a potential risk to participants which researchers must disclose when obtaining informed consent. It may affect the type of review required by some IRBs. Additionally, the IRB may require researchers to put certain protections in place, such as screening for individuals who are prone to motion sickness or conducting a MS or SS questionnaire to identify participants who are experiencing SS symptoms during the course of the study. Finally, thorough researchers should develop procedures in advance for managing situations in a compassionate and professional manner when a participant becomes sick. This may include providing sick participants with a comfortable environment, food or beverages as appropriate, as well as safely handling any biohazards (e.g. vomit) in the laboratory.

5.2.1. Informed consent

Federal Regulations regarding the protection of human subjects require researchers to obtain informed consent from each research participant and, in seeking that consent, to provide certain information, including “a description of any reasonably foreseeable risks or discomforts to the subject” (45 CFR 46). Research regarding the incidence of SS suggests that SS is a reasonably foreseeable risk that should be disclosed to research participants. Because of the many factors contributing to SS, however, researchers will not always be able to predict who will or will not get sick during simulation. Thus, it is important to disclose the risk of SS to subjects up front so that those who are prone to MS may exclude themselves from the study.

5.2.2. Minimizing risks to subjects

Not only do federal regulations require researchers to minimize risks to subjects, but doing so may actually lower the level or type of institutional ethics review. Researchers may significantly lower the level of risk involved by establishing procedures to initially exclude participants who are prone to MS and to identify participants who are demonstrating signs of becoming sick during the course of the study. As discussed throughout this article, the current research utilizes the MSAQ, administered after each driving session during the course of a study, to identify indications of SS so that participants and experimenters may stop a study before sickness occurs.

5.2.3. Dealing with SS in the lab

Despite any safeguards that researchers may put in place to prevent SS in the laboratory, at some point a participant is likely to get sick. By developing procedures in advance to deal with these situations, researchers may ensure that sick participants are adequately treated and that the lab environment is not compromised. Additionally, some IRBs may require researchers to report incidents of SS, particularly those involving serious outcomes such as vomiting. Researchers are advised to consult with their own IRB to find out what reporting mechanisms are in place at their institution and when reports are necessary.

Experimenters must be prepared for participants who become sick. Therefore, we recommend keeping sick bags, plastic gloves, mouthwash, drinks, light snacks, and cleaning products on hand. Researchers and practitioners must also consider that SS can affect performance after the simulator experience including problems with hand–eye coordination or postural instability that could interfere with the real world task of driving home (Jones et al., 2004). If an individual suffers from SS we encourage the participant to stay in the lab for a minimum of one hour and follow-up with the participant.

SS symptoms occurring some time after simulation exposure have also been reported, raising interesting questions such as how to predict when a subject might experience a delayed onset of symptoms and what researchers can do to prevent them. One of the most well known anecdotes involves a man, who after spending some time in a simulator, became so disoriented on his ride home he was forced to pull over and walk around his vehicle in order to get his bearings (Miller and Goodson, 1958). SS symptoms appearing hours after flight simulation have been well documented (Baltzley et al., 1989; Harvon and Butler, 1957; Muth, 2002). These aftereffects may present serious safety consequences for researchers and practitioners to take into consideration.

5.3. Limitations and future research

While the MSAQ may be a useful tool to determine how participants feel, it is unknown if asking these questions may cue participants into feeling ill. Future research also might investigate the relationship between responses on the MSAQ and simulator performance. Finally, we acknowledge that our use of the MSAQ

is part science and part art; thus, we have not made recommendations for specific score cutoffs to conclude that a participant's experimental session should be halted.

Acknowledgements

This work was partially supported by the Dwight D. Eisenhower Graduate Transportation Fellowship program, the National Science Foundation and the Department of Defense ASSURE Program under grant no. SES-0353698, the Creative Inquiry Initiative Program sponsored by the Office of the Provost and Undergraduate Studies at Clemson University, the South Carolina Department of Transportation and the Federal Highway Administration's School-to-Work Transportation Careers Training Program, and the Oregon Department of Transportation and the Federal Highway Administration. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funders.

We would also like to thank all of those who assisted with this project, including, Priyanka Alluri, Christine Beck, Amy DeArment, Brandon Garris, Jason Grygier, Anna Hilpert, Chris Guirl, Candice Hein, Jun Lin, Ashley Martin, Brooke Manger, Cody Palmer, Rob Ray, Kelly Riggins, Victoria Ward, and Yarbough Miller.

References

- Ardic, F.N., Topuz, B., Kara, C.O., 2006. Impact of multiple etiology on dizziness handicap. *Otology and Neurology* 27, 676–680.
- Baltzley, D.R., Kennedy, R.S., Berbaum, K.S., Lilienthal, M.G., Gower, D.W., 1989. The time course of postflight simulator sickness symptoms. *Aviation, Space, and Environmental Medicine* 60, 1043–1048.
- Bles, W., Bos, J.E., de Graaf, B., Groen, E., Wertheim, A.H., 1998. Motion sickness – only one provocative conflict? *Brain Research Bulletin* 47 (5), 481–487.
- Brooks, J.O., Tyrrell, R.A., Stephens, B.R., 2007. The accuracy of observers' estimates of their ability to see and steer in low luminances. *Journal of Vision* 7 (9), 246.
- Burke, M., 1995. Dizziness in the elderly: etiology and treatment. *The Nurse Practitioner* 20 (28), 31–35.
- Chronbach, L.J., Furby, L., 1970. How we should measure “change” – or should we? *Psychological Bulletin* 70 (1), 68–80.
- Cobb, S., Nichols, S., Ramsey, A., Wilson, J., 1999. Virtual reality-induced symptoms and effects (VRIFE). *Presence* 8 (2), 169–186.
- Drexler, J. M., 2006. Identification of system design features that affect sickness in virtual environments. Ph.D. dissertation, University of Central Florida, United States – Florida. Retrieved November 21, 2008, from Dissertations Theses: full text database (publication no. AAT 3210350).
- Duh, H.B., Parker, D.E., Furness, T.A., 2004. An independent visual background reduced simulator sickness in a driving simulator. *Presence* 13 (5), 578–588.
- Ebenholtz, S.M., 1992. Motion sickness and oculomotor systems in virtual environments. *Presence* 1 (3), 302–305.
- Ebenholtz, S.M., 2001. *Oculomotor Systems and Perception*. Cambridge University Press.
- Gianaros, P.J., Muth, E.R., Mordkoff, J.T., Levine, M.E., Stern, R.M., 2001. A questionnaire for the assessment of the multiple dimensions of motion sickness. *Aviation, Space, and Environmental Medicine* 72, 115–119.
- Guignard, J.C., McCauley, M.E., 1982. Motion sickness induced by complex periodic waveforms. *Aerospace Medicine* 53, 554.
- Harvon, M.D., Butler, L.F., 1957. Evaluation of Training Effectiveness of the 2FH2 Helicopter Flight Trainer Research Tool. Tech. Rep. No. NAVTRADEVEN 1915-00-1. Port Washington, NY: Naval Device Center.
- Hein, C., 2007. Comparison of effectiveness of roadway design treatments for transitioning from rural areas to urban areas using a driving simulator. M.S. dissertation, Clemson University, United States – South Carolina. Retrieved November 21, 2008, from Dissertations & Theses at Clemson University database (publication no. AAT 1445878).
- Hettinger, J.L., Berbaum, K.S., Kennedy, R.S., Dunlap, W.P., Nolan, M.D., 1990. Vection and simulator sickness. *Military Psychology* 2 (3), 171–181.
- Jones, M.B., Kennedy, R.S., Stanney, K.M., 2004. Toward systematic control of cyber-sickness. *Presence* 13 (5), 589–600.
- Karkos, P.D., Leong, S.C., Arya, A.K., Papouliakos, S.M., Apostolidou, M.T., Issing, W.J., 2007. Complementary ENT: a systematic review of commonly used supplements. *The Journal of Laryngology and Otology* 121, 779–782.
- Kellogg, R.S., Kennedy, R.S., Graybiel, A., 1965. Motion sickness symptomatology of labyrinthine defective and normal subjects during zero gravity maneuvers. *Aerospace Medicine* 36, 315–318.
- Kennedy, R.S., Stanney, K.M., Dunlap, W.P., 2000. Duration and exposure to virtual environments: sickness curves during and across sessions. *Presence* 9 (5), 463–472.

- Kennedy, R.S., Lane, N.E., Berbaum, K.S., Lilienthal, M.G., 1993. Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology* 3 (3), 203–220.
- Klein, N., Brooks, J., 2008. Impact of luminance and blur combinations on older drivers acuity and preferred speed. *Journal of Vision* 8 (6), 425.
- Lampton, D.R., Kolasinski, E.M., Knerr, B.W., Bliss, J.P., Bailey, J.H., Witmer, B.G., 1994. Side effects and aftereffects of immersion in virtual environments. In: *Proceedings of the Human Factors Society 38th Annual Meeting*, Human Factors Society, Santa Monica, CA, pp. 1154–1157.
- Lerman, Y., Sadovsky, G., Goldberg, E., Kedem, R., Peritz, E., Pines, A., 1993. Correlates of military tank simulator sickness. *Aviation, Space and Environmental Medicine* 64 (7), 619–622.
- Martin, A.A., Manger, B.C., Klein, N.D., Tyler, P.J., Brooks, J.O., 2007. Preferred driving speeds of older and younger drivers under varying luminance conditions. *Journal of Vision* 7 (9), 247.
- Miller, J. W., Goodson, J. E., 1958. A note concerning 'motion sickness' in the 2-FH-2 hover trainer (Project NM 1701-11, Subtask 3, Rep. 1). Pensacola, FL: U.S. Naval School of Aviation Medicine.
- Money, K.E., Myles, W.S., 1974. Heavy water nystagmus and effects of alcohol. *Nature* 247, 404–405.
- Money, K.E., Cheung, B.S., 1983. Another function of the inner ear: facilitation of the emetic response to poisons. *Aviation, Space, and Environmental Medicine* 54, 208–211.
- Muth, E.R., 2002. User considerations in virtual environments. In: *Proceedings of Human Factors and Ergonomics Society, 46th Annual Meeting*, Baltimore, MD.
- Reason, J.T., Brand, J.J., 1975. *Motion Sickness*. Academic Press.
- Reason, J.T., 1978. Motion sickness adaptation: a neural mismatch model. *Journal of the Royal Society of Medicine* 71 (11), 819–829.
- Regan, E.C., 1994. Some human factors issues in immersive virtual reality: fact and speculation. APRE report 94R027.
- Regan, C., 1995. An investigation into nausea and other side effects of head-coupled immersive virtual reality. *Virtual Reality* 1 (1), 17–32.
- Riccio, G.E., Stoffregen, T.A., 1991. An ecological theory of motion sickness and postural stability. *Ecological Psychology* 3 (3), 195–240.
- Roenker, D., Cissell, G., Ball, K., Wadley, G., Edwards, D., 2003. Speed-of-processing and driving simulator training result in improved driving performance. *Human Factors* 45 (2), 218–233.
- Treisman, M., 1977. Motion sickness: an evolutionary hypothesis. *Science* 197, 493–495.
- Von Brevern, M., Radtke, A., Lezius, F., Feldman, M., Ziese, T., Lempert, T., Neuhauser, H., 2007. Epidemiology of benign paroxysmal positional vertigo: a population based study. *Journal of Neurology, Neurosurgery, and Psychiatry* 78, 710–715.