Holistic Design Approach to Analyze Simulator Sickness in Motion-based Environments

Kevin F. Hulme, Ph.D.
NYSCEDII
Buffalo, NY
hulme@buffalo.edu

Lawrence T. Guzy, Ph.D.
SUNY Oneonta
Oneonta, NY
guzylt@oneonta.edu

Robert S. Kennedy, Ph.D.
RSK Assessments, Inc.
Orlando, FL
6kennedy@bellsouth.net

ABSTRACT

Although there are advantages to using motion-based simulators in driving, flight, marine, and entertainment contexts, one major disadvantage to their implementation is a common side effect known as simulator sickness (SS). The most common theory for the cause of SS is the discrepancy or mismatch between the illusion of motion created by the virtual world, and the artificial motion cue as detected by the inner ear. SS can come with a range of undesirable symptoms, including: headache, drowsiness, nausea, dizziness, vomiting and sweating. Unfortunately, there are conflicting theories that attempt to justify SS onset, and there is no known method for eliminating the problem. It has been postulated that approximately one in three people are susceptible, even in mild motion environments, and that approximately 80% of individuals exposed to virtual reality (VR) simulations for 20 minutes or more are more likely to report sickness symptoms. This is a continued cause for concern.

Studies have been conducted whose primary objective was to investigate the various "determinants" attributed to SS (e.g., transport delay, refresh rates, field of view, duration of exposure). Based on previous findings, it is supposed that the resultant effect is due to the **combination** of these (and various other) classes of factors. Proposed, but to date not fully realized, was a research strategy based on fractional factorial experimentation, suggesting that engineering psychology requires a **program of research**, not a miscellany of stand-alone experiments. (Simon, 1973; 1976; 1977).

Our primary goal is to specify a factorial-based, holistically-designed, human factors pilot study (N=15 participants, 33 experiments) using Orthogonal Arrays, which is a systematic method of testing the impact of various known or suspected contributing factors. Every vector or "experiment" in the array is defined as "orthogonal", which means that each is statistically independent of the others. Accordingly, the sequence of trials provides uniformly distributed coverage of the test domain. Ideally, this procedure will allow our research team to systematically isolate design factors (and interaction effects) of motion simulators that contribute most to SS. The desired outcome of our work is to develop design guidelines to minimize/mitigate SS in motion-based simulation environments.

ABOUT THE AUTHORS

Kevin F. Hulme serves as Senior Research Associate at the New York State Center for Engineering Design and Industrial Innovation (NYSCEDII). He received his Ph.D. at the University at Buffalo in Mechanical Engineering, with a specialty in design, simulation, and multidisciplinary optimization. He has a decade of experience with real-time motion-based driving simulation, including: motion platform cueing strategies and washout filtering, scene graphics development and optimization, and associated numerical methods.

Lawrence T. Guzy is an experimental psychologist with research interests in motion sickness, visual perception in challenging environments, and factors that affect safety in driving, and has collaborated on numerous research projects at the NASA Ames Research Center, Mountain View, California and Wright-Patterson Air Force Base, Dayton, Ohio. Dr. Guzy teaches courses in introductory psychology, research methods, and sensation and perception. Dr. Guzy is a recipient of the Chancellor's Award for Excellence in Teaching, and has been promoted to the rank of SUNY Distinguished Teaching Professor.

Robert S. Kennedy, Cdr., USN, Ret. is the President of RSK Assessments, Inc. He brings a tradition of hard work, strong scientific principles, and a drive to make the world better through examining human abilities as they exist and behave in this technological era of change. He has 50 years of experience in laboratory and field experimentation, 22 of which were the United States Navy where he served as an Aviation Experimental Psychologist.

Holistic Design Approach to Analyze Simulator Sickness in Motion-based Environments

Kevin F. Hulme, Ph.D.
NYSCEDII
Buffalo, NY
hulme@buffalo.edu

Lawrence T. Guzy, Ph.D.
SUNY Oneonta
Oneonta, NY
guzylt@oneonta.edu

Robert S. Kennedy, Ph.D.
RSK Assessments, Inc.
Orlando, FL
6kennedy@bellsouth.net

INTRODUCTION

A Motion Simulator (e.g., Stewart, 1965) is often implemented to simulate being inside an actual moving vehicle. The simulated motion is typically designed to be synchronized with visual displays and audio signals, which together provide an unconscious perception of movement, creating an "augmented reality" (Denne, 1994). Typical applications include simulators for ground and marine vehicles, aircraft and large equipment training, and some location-based entertainment. Despite the advantages that simulators offer, one major disadvantage to their implementation is a common side effect known as simulator sickness (**SS**), a motion sickness-like malady that can include a wide range of symptoms, including headache, sweating, vertigo, nausea, vomiting, and others (e.g., Johnson, 2005).

Unfortunately, there is no singular exact cause for SS (Kennedy & Fowlkes, 1992). There are numerous conflicting theories that attempt to justify its onset, and there is no known method for eliminating the problem outright. It has been postulated that approximately 33% are susceptible to simulator sickness, even in mild motion environments (Benson, 2002), and that approximately 80% of individuals exposed to VR simulations for 20 minutes are more likely to report sickness symptoms. (Kennedy & Stanney, 1998) This persistent problem is cause for concern, as motion-based VR is becoming more prevalent in many different applications.

Accordingly, the primary goal of our research is to perform a human factors pilot study to systematically survey the frequency and severity of SS, using Orthogonal Arrays, across a range of properties and parameters that have historically proven to contribute to the onset of SS. Findings of this design approach will undergo an analysis of variance (ANOVA). This pilot research will serve as a proof-of-concept for the overall technique, and will allow our research team to develop design guidelines to minimize and/or mitigate SS for larger, more statistically significant, and longer-term studies.

BACKGROUND AND MOTIVATION

Much of the past research in SS has focused on: a) measurement of symptoms (post-experiment), and b) stimuli encountered by VR participants during a simulation that often lead to SS. The former category includes the well-known Simulator Sickness Questionnaire (SSQ) (Kennedy et al., 1993), Nausea Rating Scales (e.g., So and Lo, 2001), and the latter category includes metrics for visual stimuli within a VR simulation, including scene complexity (i.e., "spatial frequencies") and scene movements (i.e., "navigation velocity") (So, Ho, & Lo, 2002).

A common cause of sickness within a virtual environment stems from "vection" - one's false impression of "self motion" when one is actually stationary (Dichgans and Brandt, 1978). An example of vection within the context of a motion simulator is when a driver or pilot is visualizing longitudinal motion (e.g., driving down a straight road with moderate speed), but without corresponding motion / acceleration / deceleration cues to make the participant actually "feel" like they are moving forward (or rearward).

There are three primary theories for the onset of SS: Sensory Conflict, Poison, and Postural Instability. Sensory Conflict (Reason and Brand, 1975) is based on the notion that discrepancies between the senses that provide information about the orientation and motion of our body cause a perceptual conflict. This conflict between vestibular and visual systems contradicts past experience of the subject, and manifests itself as the common symptoms of motion sickness. The Poison theory (Treisman, 1977) suggests an evolutionary premise – just as the ingestion of poison causes physiological effects involving the coordination of our sensory systems, so too do the undesired effects of prolonged periods within a virtual environment. It's almost as if the human body is misreading these unwanted effects as the ingestion of poison – these physiological effects act as an early warning system that "enhances survival" by tending to remove the contents of the stomach. Lastly, the Postural Instability theory posits

that SS is minimized when Postural Stability – one's ability to maintain the body in a stable, fixed, or balanced position -- is maximized (Riccio and Stoffregen, 1991). Hence, if the environment in which the user resides is foreign, or changes abruptly, postural control will be disrupted.

(Kennedy et al., 1997) used experience, intuition, and heuristics in an effort to decompose and document SS into its major contributing factors (which he referred to as the "Sunburst" figure), including: "Equipment Features" (e.g., visual discrepancies, field of view), "Kinematics" (both visual and motion), "Exposure Timing" (either sustained or distributed), "Individual Differences" (e.g., gender, age, state of fitness, etc.), and "Schedules" for those using a simulator (e.g., flight time, time in the simulator). More recently, a research study was conducted (Jones et al., 2004) to investigate the many "determinants" attributed to SS, with the goal of providing guidelines to limit SS to an acceptable, "specified level" for a "specified proportion of the time". This study served as a smaller scale attempt at what C.W. Simon had proposed 20 years previously – that is, a research strategy based on fractional factorial experimentation. This "Simonian" approach, proposes a systematic procedure for investigating the effect of multiple "manipulable determinants" that impact SS simultaneously. Simon argued that Engineering psychology requires a program of research, not a miscellany of stand-alone experiments (e.g., Simon, 1973; Simon, 1976; Simon, 1977).

In the case of SS, many past studies have identified a variety of (individual) contributing factors to SS. For example, common individual determinants for SS have been determined through studies on transport delays and refresh rates (DiZio and Lackner, 1997), susceptibility, vulnerability, or proneness to SS (Kennedy et al., 1990), and length of exposure/one's ability to adapt to environments that have a tendency to cause SS (Welch, 2002). It is likely that the resultant effect for most participants in a virtual environment is due to the combination of the various classes of factors suggested by Kennedy's Sunburst diagram. Accordingly, in this research, we performed a systematic human factors pilot study that investigated the multitude of factors commonly associated with the onset of SS. This type of factorial-based, holistic design approach (analogous to what Simon originally referred to as "Progressive Iteration") is, we feel, novel and vitally important to the analysis of SS – although it has been proposed, it has not (to date) been fully realized.

BROADER IMPACTS

The primary intent of the current study is to demonstrate the concept of holistic design as applied the analysis of SS, while using a limited study sample. Ideally, once this "program of research" has proven its worth in this introductory context, a much larger and more statistically significant study can be conducted.

As the motion simulator at NYSCEDII used for this study is geared toward (civilian) road vehicle simulation, so too is the current analysis of factors contributing to SS. Simulators continue to be used in a variety of meaningful studies, including driver rehabilitation (de Simone et al., 2007), driver performance (among drivers in varying age groups) (McGehee et al., 2004), and driver workload (Reimer et al., 2006) applications, to name just a few. For this reason, it is critical that the research community identify mechanisms to minimize or mitigate the onset and severity of SS-related symptoms.

Clearly however, the current research could be extended to military vehicles, specifically the Army's family of fighting vehicles: cars, trucks, tanks, and personnel carriers. The systematic design approach employed in this study is not limited to the analysis of SS in ground vehicles, and would be applicable on a similar analysis for SS in flight simulators, and thus of interest to the Air Force. Such findings would, for example, be useful for training simulators for operators of UAV's (e.g., the Predator), and to other agencies (e.g., the Navy, Border Patrol) with interests in training technology as used for simulators for other forms of remote monitoring vehicles.

Lastly, the current study will serve as a feeder for more advanced studies that attempt to look at alternative procedures for detecting, minimizing, and mitigating the onset and severity of SS. These procedures include:

- analysis of passenger facial expression for determining the onset and severity of SS (much in the same way that pilot fatigue has been analyzed using similar procedures),
- non-pharmacological countermeasures that can be introduced to minimize the severity of SS during simulation training, and
- acquisition of dynamic feedback from simulation participants during simulation training, which might allow for a more accurate assessment of the onset, duration, and severity of various contributing factors to the SS phenomenon.

ORTHOGONAL ARRAYS: IMPLEMENTATION

A vehicle simulator that integrates visual displays, motion, and sound has an inherent disadvantage: at best, it represents an abstraction of reality. For many people, the visual/vestibular system will be able to ascertain the difference (between simulation and reality), and the result will manifest itself as the onset of some degree of SS. Since the phenomenon has been analyzed, there are a number of known factors that have been shown to cause the onset of SS, pertaining to both the individual and to technology, ranging from demographic concerns, to simulator settings, to hardware limitations, to various others. Table 1 lists 16 such common contributing factors to SS onset, and an associated listing of possible level settings (continuous or discrete) for each factor.

When the expansiveness of the problem is viewed in this manner, it becomes clear that the large-scale analysis of the many potential factors associated with SS would be ideally suited for Orthogonal Arrays (OAs) (Taguchi Methods) (e.g., Hedayat, et al., 1999). OAs are a means of compressing the number of experiments that need to be conducted when the goal is to measure the contribution of each of several factors to a prescribed (and measurable) outcome. For example, suppose we had seven "on/off" factors in a particular experiment. Simplistically, we would need to conduct 2^7, or 128 experiments, one for each possible combination of factors, for a full factorial experiment. Alternatively, we could employ an L8 OA, as shown in Figure 1. This array involves a total of eight experiments (and contains 8 rows), with each one investigating seven factors (one per column), each assigned at one of two discrete levels.

OA's are a systematic, statistical means for testing the impact of a factor (or various factors) on a system or phenomenon. Every row and column (i.e., vector) in such an array is defined as "orthogonal", which implies that each of the vectors is statistically independent of the others, and thus when linearly combined, the resultant is the arithmetic sum of the individual components. Accordingly, the sequence of trials is efficient; more specifically, it provides uniformly distributed coverage of the test domain, represents a concise test set with fewer test cases, and thus dramatically reduces overall testing cycle time. By taking advantage of the principals of orthogonality, in the example case of the L8 array, we can systematically capture a significant portion of the solution space with just eight experiments, instead of the 128 that would be required for full factorial examination. Note that the L8 array (and all official OAs, for that matter) are generated systematically, and are typically culled by the user community from finite published lists/collections (e.g., Kuhfeld, 2011) that have been proven and validated. In other words, one cannot generate an array based on an arbitrary number of variables and experiments, and claim it to be orthogonal.

Table 1. Experimental Factors Associated with a Motion-Based Vehicle Simulation

Factor	Level Ranges
Subject Gender	M or F
Motion - Rotation DOF's	ON or OFF
Motion - Translation DOF's	ON or OFF
Drive mode	Driver or Passenger
Audio cues	ON or OFF
Time of day	Day or Night
Forward field of view	60 – 180 deg.
DOF Scaling	0 – 100%
Input/output latency	0 – 1 sec.
Hours since last meal	0-6
Laboratory Lighting	On, Off, Dim, …
Projector Brightness	0 – 100% (lumens)
DOF rate limiting	"slow" – "fast"
Exposure duration	1 – 30 mins.
Motion DOF's	0-6
Visual frame rate	15-120 Hz.
("DOF" = Degree of Freedom)	

Experiment	*Factor 1*	*Factor 2*	*Factor 3*	*Factor 4*	*Factor 5*	*Factor 6*	*Factor 7*
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2

Figure 1 – L8 OA
(eight experiments, seven 2-level variables)

Each of these experiments will have an associated observation value (i.e., "objective function") that will serve as a quantified measure of SS. For example, Kennedy's SSQ will be implemented after each experiment. These quantitative measures will subsequently allow us to determine the effect of each factor level, based on its deviation from the overall mean effect. Then, an Analysis of Variance

(ANOVA) (Scheffé, 1959) can be performed – this statistical analysis can compare mean values (of factor level effects versus the overall mean) to determine the relative importance of factors (and their levels) to the observation value, thus predicting "optimal" factor/level settings so as to minimize SS for a given set of parameters. After the "optimal" assignments of factor levels are determined to meet a design objective (i.e., "minimize SS"), a verification experiment (or experiments) can be performed to ensure the validity of the result. Figure 2 depicts a sample L8 array in the context of a prospective SS experiment, using the factors/levels seen in Table 1.

Experiment	Gender	Rotation DOF's	Translation DOF's	Driver/Passenger	Audio Cues	Day (D) or Night (N)	Forward F.O.V.
1	M	on	on	D	on	D	60
2	M	on	on	P	off	N	180
3	M	off	off	D	on	N	180
4	M	off	off	P	off	D	60
5	F	on	off	D	off	D	180
6	F	on	off	P	on	N	60
7	F	off	on	D	off	N	60
8	F	off	on	P	on	D	180

Figure 2 - Experimental L8 OA

Our human factors pilot study serves as a demonstration of a "program of research" - whose ultimate goal is to systematically attain a statistically-significant understanding of the various factors that contribute to SS, and their degree of impact.

PILOT STUDY DESCRIPTION

Facilities

The simulations for the pilot study are conducted in NYSCEDII's motion simulation laboratory, which is equipped with a passenger cabin mounted on 6-DOF motion platform, a four-channel surround-screen visualization environment, and a 2.1 channel stereo sound system. Refer to Figure 3.

Recruitment methods, Study size, and Compliance

Fifteen participants were recruited for this study, of varying age, gender, and years driving experience. The minimum age for the study is 18 years. All participants were screened for physical conditions that would prevent participation, including epilepsy, past episodes with seizures, and proneness to extreme motion sickness (i.e., air, car, or sea). People with severe phobias that related to the environment of our simulator (e.g. claustrophobia, nyctophobia) were not allowed to participate. Due to their proneness to morning sickness, pregnant women were excluded.

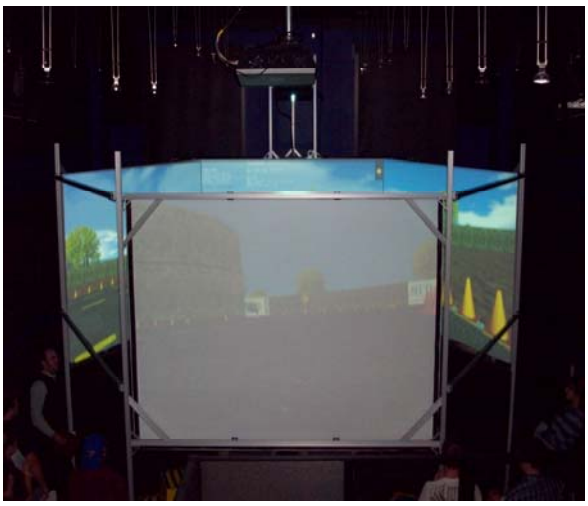

Figure 3 – NYSCEDII Motion Simulator

Note again that this is a pilot study intended to serve as a proof of concept for a much larger "program of research". With our baseline sample size (N=33 experiments, and N=15 participants), statistical significance is only marginal: (95% confidence level; 25.3% confidence interval) (CRS, 2010). A substantially larger sample size (e.g., N=1000) (95% confidence level; 3.1% confidence interval) would require a great deal more time and resources than were currently available for this project. Our sample size was chosen primarily based upon the sizes of the three chosen OAs that were implemented (and fully populated) for this paper.

The study was pre-approved by the Institutional Review Board (IRB), and was found to be in full compliance. Accordingly, human subject privacy, confidentiality, and safety were ensured at all times. Subjects were minimally compensated for their time during the 60-minute experiments.

Driving tasks

The pre-existing simulation software environment that was used for the study is an approximately a two square mile region that has been modeled after an actual array of neighborhoods, streets, and landmarks adjacent to the University at Buffalo, shown in map view in Figure 4. The environment includes

residential streets, slightly busier 2-lane roads, 4-lane roads, and a 1-mile stretch of the New York State Thruway. Our virtual driving environment is bordered by North Bailey Avenue on the West, Maple Road on the North, Sheridan Drive on the South, and both New York State Route 290, and a small section of Harlem Road on the East.

Participants are asked to drive a pre-specified road course (the black path shown in Figure 4), whose total length is approximately eight miles long, and under normal driving conditions, requires approximately 20 minutes to complete. Within Figure 4, the average 5, 10, 15, and 20-minute end points are shown, which demonstrates approximately how far each subject will drive for various pre-set durations. There are five stop signs along the course (shown as red dots) and eleven traffic lights (shown as green dots). Depending on the assigned level for the "Duration" factor, drivers were asked to complete all, a portion of this road course.

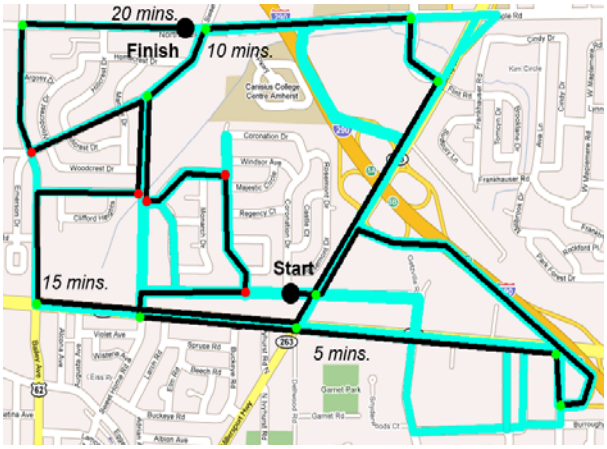

Figure 4 – SS driving environment

Session details and Data Measures

Each participant session lasted about 60 minutes, with the time divided as shown in Table 2. Depending on the assigned duration of each experiment, each participant might participate in one long experiment in the simulator (e.g., 30 minutes), or a series of shorter experiments (e.g., 3 10-minute experiments). Note that "duration" is one of the primary factors associated with SS (see Table 1), and was thus varied for certain participants and experiments. In addition to quantitative and observational driver performance data, outcome measures will include pre- and post-experiment surveys, including the SSQ the more recent Motion Sickness Assessment Questionnaire (MSAQ) (Gianaros, 2001).

Table 2 – Activity breakdown of each session

Activity	Duration
Informed Consent	10 minutes
Pre- Surveys (SSQ, MSAQ)	5 minutes
Simulator experiment(s)	30 minutes (max)
Participant rest	5 minutes
Post- Surveys (SSQ, MSAQ)	5 minutes
Participant observation	5 minutes
TOTAL:	**60 minutes**

RESULTS AND DISCUSSION

Chosen OA's for experiments

To demonstrate the concept of using OAs in this context on a small scale, three different arrays were selected: one each with 2-level factors (L8), 3-level factors (L9), and 4-level factors (L16). This amounts to a total of 33 driving experiments, for which a study size of N=15 participants were selected. Our L8 array can be seen in Figure 2, and our L9 and L16 arrays can be seen in Figures 5 and 6, respectively.

Experiment	***DOF Scaling***	***I/O latency***	***Hours last meal***	***Laboratory Lighting***
1	20%	0	<2	ON
2	20%	0.125	2-4	OFF
3	20%	0.25	>4	Dim
4	40%	0	2-4	Dim
5	40%	0.125	>4	ON
6	40%	0.25	<2	OFF
7	60%	0	>4	OFF
8	60%	0.125	<2	Dim
9	60%	0.25	2-4	ON

Figure 5 – Experimental L9 OA

Table 3 summarizes a variety of general details regarding our study population.

Table 3 – Summary of study population

Population size	15
Average age	21.2 (±1.9)
Males	10
Females	5
Average years driving	3.7 (±2.5)

Results and Analysis

Two primary metrics were employed for the study: the SSQ and the MSAQ. The MSAQ is normalized on a 0 (not sick) to 1 scale, and the SSQ scale ranges from 0 (not sick) to a maximum possible score of approximately 235. Further details for each technique can be found in the literature.

Experiment	Scene Brightness	DOF rate limiting	Exposure duration	Motion DOF's	Visual frame rate
1	15%	0.001	5	0	15
2	15%	0.005	10	2	24
3	15%	0.0075	15	4	45
4	15%	0.01	20	6	60
5	40%	0.001	10	4	60
6	40%	0.005	5	6	45
7	40%	0.0075	20	0	24
8	40%	0.01	15	2	15
9	65%	0.001	15	6	24
10	65%	0.005	20	4	15
11	65%	0.0075	5	2	60
12	65%	0.01	10	0	45
13	90%	0.001	20	2	45
14	90%	0.005	15	0	60
15	90%	0.0075	10	6	15
16	90%	0.01	5	4	24

Figure 6 – Experimental L16 OA

Mean (overall) scores for the study, for each of our two outcome measures (the SSQ and the MSAQ), are listed in Table 4.

Table 4 – Mean (overall) Scores

	SSQ	MSAQ
L8	52.87	0.185
L9	28.05	0.154
L16	41.06	0.208

Table 5 displays the factor effects for the L8 array, using the SSQ rating measure. Note that the "optimal" (smaller) level setting for each factor has been underlined in the Table. Figures 7 and 8 illustrate the factor effects for the L9 and L16 arrays, respectively, using the MSAQ rating measure.

Table 5 – Main effects for L8 array (SSQ)

Factor (setting 1, setting 2)	1	2
Gender (Male, Female)	<u>35.48</u>	70.26
Rotation DOF (On, Off)	<u>36.60</u>	69.14
Translation DOF (On, Off)	71.80	<u>33.94</u>
Ride Mode (Driver, Passenger)	55.81	<u>49.92</u>
Audio cues (On, Off)	56.24	<u>49.50</u>
Time-of-day (Day, Night)	58.76	<u>46.98</u>
Field of View (FOV) (60^o, 180^o)	<u>52.45</u>	53.29

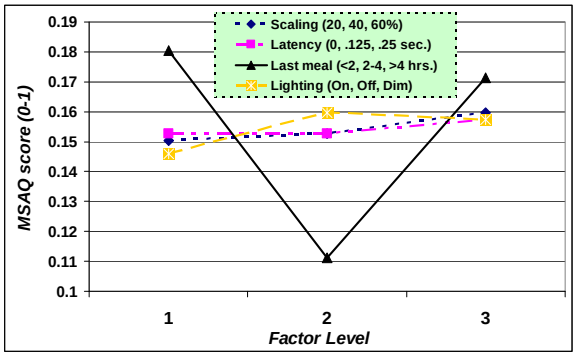

Figure 7 – Main factor effects (L9 array)

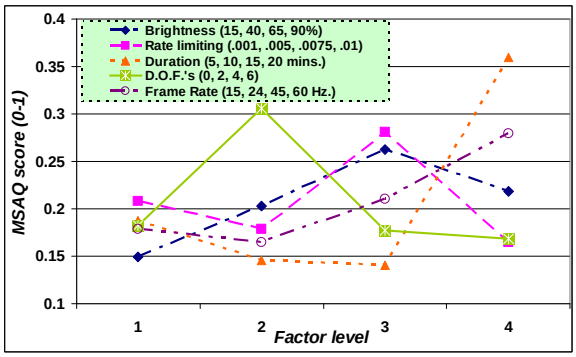

Figure 8 – Main factor effects (L16 array)

As part of our ANOVA, we also calculated the sum of squares due to each factor, which provides the relative importance of each factor in varying the overall mean. Once again, for the L8 array, we utilize the SSQ measure, and for the L9 and L16 arrays, we utilize the MSAQ measure. Grand total sum of squares (GTSOS), sum of squares due to mean (SOSM), and the total sum of squares (TSOS) for each OA are offered in Table 6, and Figures 9-11 plot the sum of squares for each factor, for the L8, L9, and L16 arrays, respectively.

Table 6 – ANOVA (sum of squares)

	L8 (SSQ)	L9 (MSAQ)	L16 (MSAQ)
GTSOS	30209.44	0.223	0.903
SOSM	22365.49	0.214	0.694
TSOS	7843.94	0.009	0.209

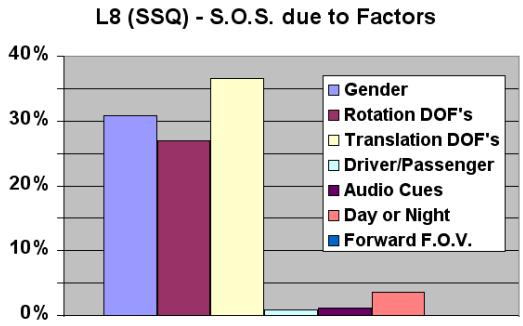

Figure 9 – sum of squares (SSQ, L8 array)

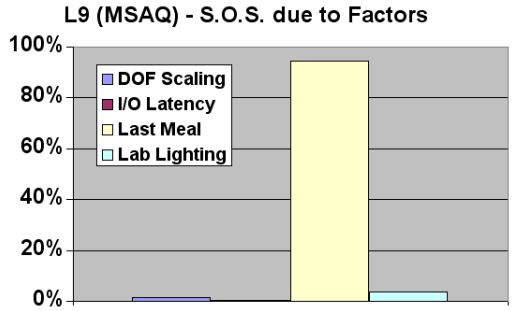

Figure 10 – sum of squares (MSAQ, L9 array)

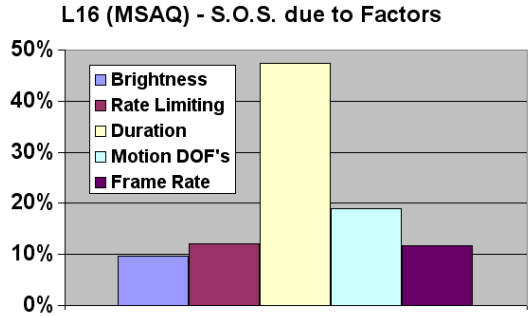

Figure 11 – sum of squares (MSAQ, L16 array)

Discussion

A primary goal of this study was to demonstrate the technique of using OAs to systematically analyze multiple factors associated with SS simultaneously. A secondary goal of the study was to identify patterns in the data that correspond with findings in past literature. Accordingly, once the experiments were complete, our objective was to interpret the trends of the main effects of the SS factors, and their various levels. For this study, our N=15 subject, 33-experiment sample size proved too small to perform anything more than a basic ANOVA. Accordingly, we investigated "main effects", and decided to save interaction effects for a follow-up study.

During the planning phases for this study, it was decided not to take the N=15 study population "to the limit" (i.e., vomiting, or other extreme sickness symptoms) during testing. Rather, we decided it was more appropriate for subjects to self-rate themselves after each driving experiment based upon their perception of their own symptoms. Accordingly, our mean SSQ and MSAQ scores (Table 4) were not insignificant, but were on the low side of the scale, with the highest mean ratings somewhere around 20% of the maximum value for each scale.

Reviewing the L8 array results shown in Table 5, and comparing these values to the SSQ listed mean scores shown in Table 4, we can make the following observations. The females who participated in our study were much more susceptible to SS than the males. Rotation cues (roll, pitch yaw), as expected, minimized SS effects, while Translation cues (heave, surge, sway) seemed to have an adverse effect. The latter is either because the sample size was too small, or because our washout algorithm cues need to be more properly tuned. Also surprising is the fact that vehicle passengers seemed to be less prone to SS that drivers. This could also be due to the sample size, or more likely, is due to the fact that passengers could relax and enjoy the simulation without the pressure of having to "perform" as a driver. Almost all subjects entered our Laboratory with great anticipation and anxiety, so enjoying the simulator as a passenger seemed to have a relaxing effect on many. Audio cues turned ON seemed to induce slightly more SS, which again is surprising for the reason that drivers seemed to drive much more aggressively WITHOUT sound, as they had no aural feedback regarding how fast they were going (i.e. the increasing pitch/gain of the engine idle). One would have expected that this aggressive type driving (e.g., fast starts and abrupt stops) would have resulted in more pronounced SS symptoms, but the data for this limited sample did not back this up. Driving at night (as opposed to day) seemed to lessen SS symptoms, which was expected, as dimmer images historically seem to be more pleasant to look at within a virtual environment. Finally, forward FOV seemed to have little effect, which is again surprising, as one would expect a wider FOV to be more authentic and immersive, but also more likely to induce a stronger sense of vection.

Reviewing the L9 array results shown in Figure 7, DOF scaling seemed to induce more SS with greater percentages. This is indicative that some supplementary motion is a good thing, but too much can be detrimental. This result is expected, but we would expect that the "optimal" value for DOF scaling would be greater than 20%, the smallest value examined in this study. I/O latency did increase SS symptoms proportionately, but seemed to have very

little overall effect, which is somewhat surprising. Perhaps our novice pool of drivers didn't notice the small/nonzero latencies; it may have been worthwhile to introduce larger and more obvious latencies to enhance the effect, and overall impact. Eating a meal before experiencing the simulator seemed to be optimal at 2-4 hours prior to the experiment, as opposed to less than two hours before, or more than four hours before. Much more data is needed to support such findings, but it does seem likely that one would be most comfortable on a simulator having eaten (and raised ones blood sugar level) not too long, yet not too soon prior to participating. Finally, Laboratory Lighting seemed to induce the smallest SS effects while in the ON position, as opposed to OFF or Dim. Perhaps the subjects felt less immersed and more relaxed and "detached" from the simulation environment with the house lights turned on.

Reviewing the L16 results shown in Figure 8, as expected, a dimly lit scene induced the smallest level of SS, and the general trend was that as projector brightness increases, so too do the observed effects of SS. DOF rate limiting was found to be most desirable at the highest value (0.01), meaning, the motion platform was most responsive at this value. Not surpassing was the dramatic increase in observed SS symptoms for the long duration trials (20 minutes), although the trend was not found to be linear, as the second largest level effect for the duration factor, after twenty minutes, was the shortest duration (of just five minutes). This anomaly is likely attributable to the small sample size. Also not surprising was the fact that full motion (6 DOF's) was found to be the most comfortable scenario, but a curious spike in observed SS symptoms occurred with the 2-DOF (roll/pitch) scenario. These are the two most prominent DOF's, but perhaps felt foreign in the absence of the other four DOF's. We suspect that an appropriate sample size that show that the number of DOF's utilized is inversely proportional to the amount of undesirable SS effects encountered. Finally, frame rate was found to be most desirable at 24 Hz., and least desirable at 60 Hz, which was surprising, and likely (again) due to insufficient size of our sample. While screen flicker was very noticeable for 15 Hz, not detectable for 45 and 60 Hz., and perhaps slightly detectable for the 24 Hz. setting, we would expect that frame rate is also inversely proportional to observed SS symptoms.

Finally, Figures 9-11 illustrate the relative importance of the chosen SS factors for our three experimental arrays. In the L8 array (Figure 9), "gender", and "rotation/translation DOF's" were found to be dominant, with day/night cues having some contribution. In the L9 array (Figure 10), "hours since last meal" was by far the dominating factor. And finally, in the L16 array (Figure 11), there was much more balance between all 5 factors present, with "duration" being the dominant factor.

FUTURE WORK

Expanded Pilot Study

A logical next step would be to perform a larger-scale study, where many SS contributing factors could be analyzed within the confines of a single OA. Accordingly, one might select the L96 array (e.g., Kuhfeld, 2011), which can accommodate 39 2-level factors, 1 3-level factor, 14 4-level factors, and 1 8-level factor. One (or more) columns could then be reserved for studying interaction effects among factors; not considered in the present study.

Facial Expression Recognition

For pilots aboard flight simulators, it might be useful to have a non-invasive mechanism to detect/predict the onset of SS without the need for physiological monitoring hardware (e.g., blood pressure, respiration, perspiration). Research has shown that when emotions are elicited, fleeting traces of its corresponding facial expression are expressed by unconscious movements of facial muscles, despite conscious efforts to conceal the expression (Frank and Ekman, 2004). These would serve as reliable indicators for facial patterns linked to symptoms commonly associated with SS (e.g., nausea).

Countermeasures

Recent research (Guzy et al., 2008) has demonstrated promise for incorporating non-pharmacological mechanisms such as vibro-tactile stimulation and 3-D stereo sound to improve situational awareness within a virtual environment, and ideally to minimize any associated undesirable effects. Another recent study (Fleur et al., 2006) attempted behavioral methods (e.g. controlling breathing) in an effort to reduce sickness symptoms, and increasing tolerance to motion-induced nausea.

Dynamic Feedback

We will experiment dynamically with experimental factors to determine the threshold at which SS symptoms begin, or heighten. During a simulation, we will alter a parameter that might induce SS, and then every 30 seconds, request feedback from the simulation participant. The participant will then convey their perceived level of SS on a 0-10 scale. Participant answers will be recorded/logged, and their feedback will be compared to the corresponding experimental factor settings at that timestamp.

SUMMARY AND CONCLUSIONS

Despite the numerous advantages that motion simulators offer, one major disadvantage to their implementation is simulator sickness (**SS**). Approximately 33% are susceptible to simulator sickness, even in mild motion environments, and 80% of individuals exposed to VR simulations for 20 minutes are more likely to report sickness symptoms.

In this research, we performed a human factors pilot study to survey the frequency and severity of SS using OAs, which compress the number of experiments that need to be conducted when the goal is to measure the contribution of each of several factors to a measurable outcome. Every row and column in an OA is statistically independent of the others. Accordingly, the sequence of trials is efficient, and systematically captures a significant portion of the solution space without full factorial examination. This technique serves as a "program of research", as opposed to more common approaches that employ a sequence of stand-alone experiments.

Fifteen participants were recruited to perform a total of 33 experiments, which were used to populate the L8, L9 and L16 OAs, which had seven, four, and five factors, and two, three, and four levels-per-factor, respectively. Participants were asked to drive the motion simulator along a pre-defined road course of varying durations, while simulation parameters were varied for each trial. Measures included the SSQ and the MSAQ, and were offered after each driving trial.

A primary goal of this study was achieved, which was to demonstrate the technique of using OAs to systematically analyze multiple factors associated with SS simultaneously. The study was able to recognize various patterns in the data that correspond with findings in past literature, and the authors were able to interpret, and to some degree, justify the observed trends of the main effects of the SS factors, and their various levels.

This pilot research serves as a proof-of-concept for larger, more statistically significant, and longer-term studies that consider all factors/levels simultaneously as well as interaction effects. Our team likewise identified numerous avenues for future work in recognizing and mitigating SS onset and severity.

ACKNOWLEDGEMENTS

The authors would like to acknowledge hardware and software support from Moog, Inc., subject recruitment assistance from Ms. Karen Morris at the University at Buffalo, and funding support from the State of New York, and NYSTAR.

REFERENCES

Benson, A.J. (2002). "35. Motion Sickness". Medical Aspects of Harsh Environments, Volume II. Borden Institute, Washington, DC.

Creative Research Systems (CRS), (2010). "Sample Size Calculator", (online web link), http://www.surveysystem.com/sscalc.htm, Copyright 2007-2010.

Denne, P., (1994). "Motion Systems and Visual Displays", Unpublished White Paper, January, 1994.

de Simone, V., Kaplan, L., Petronas, N., Wassermann, E.M., and Grafman, J., (2007). "Driving Abilities in Frontotemporal Dementia Patients," *Dementia and Geriatric Cognitive Disorders*, 23, pp. 1–7, DOI: 10.1159/000096317.

Dichgans, J., and Brandt, T., (1978). "Visual-Vestibular Interaction: Effects of Self-Motion Perception and Postural Control", in R. Held, H. W. Leibowitz and H. L. Teuber (Eds.), Handbook of Sensory Psychology, Volume VIII: Perception., Berlin: Springer-Verlag.

DiZio, P., and Lackner, J. R.. (1997). "Circumventing side effects of immersive virtual environments", In M. Smith & R. Salvendy (Eds.), Design of computing systems: Social and ergonomic considerations, pp. 893-896, Amsterdam: Elsevier.

Fleur, D., Yen Pik, S., Billar, J.P., Golding, J.F., and Gresty, M.A., (2006). "Behavioral Methods of Alleviating Motion Sickness: Effectiveness of Controlled Breathing and a Music Audiotape", *Journal of Travel Medicine*, Vol. 10, No.2, pp. 108-111.

Frank, M.G., and Ekman, P. (2004). "Appearing truthful generalizes across different deception situations", *Journal of Personality and Social Psychology*, 86, 486-495.

Gianaros, P.J., Muth, E.R., Mordkoff, J.T., Levine, M.E., and Stern, R.M., (2010). "A Questionnaire for the Assessment of the Multiple Dimensions of Motion Sickness", *Aviat Space Environ Med*. 2001 February; 72(2): 115–119.

Guzy, L.T., Albery, W.B., and Goodyear, C., (2008). "Vibrotactile stimulators and virtual 3-D audio countermeasures, training and motion sickness symptoms with a simulated graveyard spin illusion", *Journal of Vestibular Research*, Vol. 18, No. 5-6, pp. 287-294.

Hedayat, A.S., Sloane, N.J.A., and Stufken, J., (1999). Orthogonal Arrays: Theory and Applications, Springer-Verlag, New York.

Johnson, D.M., (2005). "Introduction to and Review of Simulator Sickness Research", U.S. Army Research Institute for the Behavioral and Social Sciences, Research Report 1832, April 2005.

Jones, M.B., Kennedy, R.S., and Stanney, K.M., (2004). "Toward systematic control of Cybersickness", *Presence: Teleoperators and Virtual Environments*, Vol. 13, No.5, pp. 589-600, ISSN:1054-7460.

Kennedy, R.S., Dunlap, W.P., and Fowlkes, J.E., (1990). "Prediction of motion sickness susceptibility", In G.H. Crampton (Ed.), Motion and space sickness, pp.179-215, Boca Raton, FL: CRC Press.

Kennedy, R.S., and Fowlkes, J.E., (1992). "Simulator Sickness is Polygenic and Polysymtomatic: Implications for Research", *International Journal of Aviation Psychology*, 2(1): 23-38.

Kennedy, R.S., Lane, N.E., Berbaum, K.S., and Lilienthal, M.G., (1993). "Simulator sickness questionnaire: An Enhanced Method for Quantifying Simulator Sickness", *The International Journal of Aviation Psychology*, Vol.3, No.3, pp. 203-220.

Kennedy, R.S., Drexler, J.M., and Compton, D.E., (1997). "Simulator Sickness and Other Aftereffects: Implications for the Design of Driving Simulators", Proceedings of the Driving Simulation Conference - DSC'97, ETNA, Paris, France, pp. 115-123.

Kennedy, R.S., and Stanney, K.M., (1998). "Aftereffects from virtual environment exposure: How long do they last?", Proceedings of the 42nd annual meeting of the Human Factors and Ergonomics Society, pp.1476-1480.

Kuhfeld, W.F., (2011). "Orthogonal Arrays", SAS online web link, http://support.sas.com/techsup/technote/ts723.html, Copyright 2011.

McGehee, D.V., Lee, J.D., Rizzo, M., Dawson, J., and Bateman, K., (2004). "Quantitative Analysis of Steering Adaptation on a High Performance Fixed-base Driving Simulator", Transportation Research Part F: Traffic Psychology and Behavior, 7(3), pp. 181-196.

Reason, J.T., and Brand, J.J., (1975). "Motion Sickness", London: Academic Press.

Reimer, B., Mehler, B.L., Pohlmeyer, A.E., Coughlin, J.F. and Dusek, J.A. (2006). "The Use of Heart Rate in a Driving Simulator as an Indicator of Age-Related Differences in Driver Workload," *Advances in Transportation Studies - an International Journal*, pp. 9-20.

Riccio, G.E., and Stoffregen, T.A., (1991). "An Ecological Theory of Motion Sickness and Postural Instability", *Ecological Psychology*, 3(3):195-240.

Scheffé, H. (1959). The Analysis of Variance, New York: John Wiley & Sons.

Simon, C.W., (1973). "Economical multifactor designs for human factors engineering experiments", Technical Report No. P73-326A. Culver City, CA: Hughes Aircraft Company.

Simon, C.W., (1976). "Analysis of Human Factors Engineering Experiments: Characteristics, Results, and Applications", Technical Report No. CWS-02-76, Westlake Village, CA: Canyon Research Group.

Simon, C.W., (1977). "New research paradigm for applied experimental psychology: a system approach", Technical Report No. CWS-04-77A, Westlake Village, CA: Canyon Research Group, Inc.

So, R.H.Y, and Lo, W.T., (2001). "Cybersickness in the Presence of Scene Rotational Movements Along Different Axes", *Applied Ergonomics*, Vol.32, pp.1-14.

So, R.H.Y, Ho, A.T.K., and Lo, W.T., (2002). "A Metric to Quantify Virtual Scene Movement for the Study of Cybersickness: Definition, Implementation, and Verification", *Presence*, Vol.10, No.2, pp.193-215.

Stewart, D., (1965). "A platform with six degrees of freedom", *The Institution of Mechanical Engineers*, 180(15): pp. 371-384.

Treisman, M., (1977). "Motion Sickness: An Evolutionary Hypothesis", *Science*, Vol. 197, pp. 493-495.

Welch, R.B., (2002). "Adapting to virtual environments", In K.M. Stanney (Ed.), Handbook of virtual environments: Design, implementation, and applications, pp. 619-636, Mahwah, NJ: Lawrence Erlbaum Associates, Inc.