Evaluation of Virtual Reality Snowplow Simulator Training: Final Report

Final Report
January 2007

Sponsored by
the Iowa Department of Transportation
(CTRE Project 06-245)

Center for Transportation Research and Education

Iowa State University

Iowa State University's Center for Transportation Research and Education is the umbrella organization for the following centers and programs: Bridge Engineering Center • Center for Weather Impacts on Mobility and Safety • Construction Management & Technology • Iowa Local Technical Assistance Program • Iowa Traffic Safety Data Service • Midwest Transportation Consortium • National Concrete Pavement Technology Center • Partnership for Geotechnical Advancement • Roadway Infrastructure Management and Operations Systems • Statewide Urban Design and Specifications • Traffic Safety and Operations

About CTRE/ISU

Disclaimer Notice

Non-discrimination Statement

Technical Report Documentation Page

1. Report No. CTRE Project 06-245	**2. Government Accession No.**	**3. Recipient's Catalog No.**	
4. Title and Subtitle Evaluation of Virtual Reality Snowplow Simulator Training: Final Report		**5. Report Date** January 2007	
		6. Performing Organization Code	
7. Author(s) Christopher M. Masciocchi, Veronica J. Dark, Derrick Parkhurst		**8. Performing Organization Report No.**	
9. Performing Organization Name and Address Center for Transportation Research and Education Iowa State University 2711 South Loop Drive, Suite 4700 Ames, IA 50010-8664		**10. Work Unit No. (TRAIS)**	
		11. Contract or Grant No.	
12. Sponsoring Organization Name and Address Iowa Department of Transportation 800 Lincoln Way Ames, IA 50010		**13. Type of Report and Period Covered** Final Report	
		14. Sponsoring Agency Code	
15. Supplementary Notes Visit www.ctre.iastate.edu for color PDF files of this and other research reports.			
16. Abstract Each winter, Iowa Department of Transportation (Iowa DOT) maintenance operators are responsible for plowing snow off federal and state roads in Iowa. Drivers typically work long shifts under treacherous conditions. In addition to properly navigating the vehicle, drivers are required to operate several plowing mechanisms simultaneously, such as plow controls and salt spreaders. There is little opportunity for practicing these skills in real-world situations. A virtual reality training program would provide operators with the opportunity to practice these skills under realistic yet safe conditions, as well as provide basic training to novice or less-experienced operators. In order to provide such training to snowplow operators in Iowa, the Iowa DOT purchased a snowplow simulator. The Iowa DOT commissioned a study through Iowa State University designed to (1) assess the use of this simulator as a training tool and (2) examine personality and other characteristics associated with being an experienced snowplow operator. The results of this study suggest that Iowa DOT operators of all ages and levels of experience enjoyed and seemed to benefit from virtual reality snowplow simulator training. Simulator sickness ratings were relatively low, implying that the simulator is appropriate for training a wide range of Iowa DOT operators. Many reported that simulator training was the most useful aspect of training for them.			
17. Key Words driving simulators—eye movement—job performance—snowplow training—virtual reality		**18. Distribution Statement** No restrictions.	
19. Security Classification (of this report) Unclassified.	**20. Security Classification (of this page)** Unclassified.	**21. No. of Pages** 75	**22. Price** NA

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized**

EVALUATION OF VIRTUAL REALITY SNOWPLOW SIMULATOR TRAINING: FINAL REPORT

Final Report
January 2007

Principal Investigator
Derrick Parkhurst
Associate Director, Virtual Reality Application Center
Assistant Professor, Department of Psychology
Iowa State University

Co-Principal Investigator
Veronica J. Dark
Associate Professor, Department of Psychology
Iowa State University

Research Assistant
Christopher M. Masciocchi

Authors
Christopher M. Masciocchi, Veronica J. Dark, Derrick Parkhurst

Preparation of this report was financed in part
through funds provided by the Iowa Department of Transportation
through its research management agreement with the
Center for Transportation Research and Education,
CTRE Project 06-245.

A report from
Center for Transportation Research and Education
Iowa State University
2711 South Loop Drive, Suite 4700
Ames, IA 50010-8664
Phone: 515-294-8103
Fax: 515-294-0467
www.ctre.iastate.edu

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

The authors would like to thank the Iowa Department of Transportation for sponsoring this research.

INTRODUCTION

Overview

Each winter, Iowa Department of Transportation (Iowa DOT) maintenance operators are responsible for plowing snow off federal and state roads in Iowa. Drivers typically work long shifts under treacherous conditions. In addition to properly navigating the vehicle, drivers are required to operate several plowing mechanisms simultaneously, such as plow controls and salt spreaders. There is little opportunity for practicing these skills in real-world situations. During snowfalls, when training would be most realistic and effective, all available vehicles, drivers, and potential instructors are required to be plowing on the roadways. Consequently, novice operators often do not undergo as comprehensive a training regimen as desired, and experienced operators do not have opportunities to improve their current practices or test new ones. Additionally, conducting novice snowplow operator training on roadways may present an unnecessary hazard for the trainee as well as other drivers.

Virtual reality training is an option when real-world training would be prohibitively high-priced, inappropriate, or hazardous. For example, because they provide a safe yet realistic environment in which students can be taught how to operate aircraft, practice in flight simulators is a basic requirement of pilot training programs. A similar training program for snowplow operators would provide experienced operators with a chance to practice their skills under realistic yet safe conditions and would supply basic training to novice or less-experienced operators. Training could be conducted during any time of year, which would be especially beneficial during the summer when vehicles and drivers are not pre-engaged. In order to provide such training to snowplow operators in Iowa, the Iowa DOT purchased a snowplow simulator from L3 Communications in 2005.

The Iowa DOT commissioned a study through Iowa State University designed to (1) assess the use of this simulator as a training tool and (2) examine personality and other characteristics associated with being an experienced snowplow operator. This manuscript is the technical report of that study. The manuscript begins with a brief and general review of the simulation literature before proceeding to a detailed description of the study. A more extensive literature review designed to complement this study is available in Masciocchi, Dark, and Parkhurst (2006).

The detailed description of the study begins with the method and then addresses the following questions:

- What are the personality and demographic characteristics of Iowa DOT snowplow operators?
- How do participants respond to and evaluate the simulator training?
- Do the effects of training vary with experience?
- Does training increase the ability of operators to successfully and efficiently complete simulation scenarios?

Review of the Simulation Literature

As already suggested, simulators are used to provide a safe and inexpensive practice environment for persons who operate complex machines (e.g., airplanes, trucks, cars, and snowplows) in the real world. Recent technological advances, such as those allowing multiple screens to increase the field of view, have improved the fidelity of the virtual-reality environments in simulators, and thus increased their applicability in training or research settings. Kemey and Panerai (2003), for instance, concluded that driving simulators with a wide field of view, more than 120 degrees, provided optic flow and thus were able to provide drivers with a sense of motion. Fast screen refresh rates, above the 30Hz rate at which humans detect screen flicker, also contribute to realism and reduce the likelihood of cybersickness (LaViola 2000). Several studies have found similarities in participants' behavior in realistic driving simulators and their behavior in real vehicles (Godley, Triggs, and Fildes 2002; Panerai et al. 2001; Tornros 1998). For example, Panerai et al. and Tornros found that participants drove approximately the same speed in simulator and real-world conditions when their speedometers were covered, and Tonros reported that participants in driving simulators and real vehicles altered their speed similarly based on cues in the environment.

Since behaviors in simulators and in actual driving situations are correlated, driving simulators are now commonly used to investigate how drivers react to real-life situations that would be difficult or costly to replicate in non–virtual reality environments. For instance, Horberry, Anderson, and Regan (2005) examined the effects of standard versus enhanced road markings on driving behavior. Because previous research had suggested that the presence of lane markings reduces accidents (Miller 1992), Horberry et al. (2005) reasoned that making lane markings more prominent, especially in low visibility conditions, might lead to a further reduction in accidents. Their results showed that participants in the enhanced marking condition drove closest to the target speed, showed less variability in their lane position, were less likely to cross over the center line at inappropriate times, and rated the roads with enhanced marking as easier to drive on. The results confirmed the predictions of experts and the observations of real-world driving studies showing that increasing the prominence of lane markings would positively affect drivers' performance under poor visibility conditions. The study by Horberry et al. illustrates the role that simulators have come to play in real-world tasks such as designing road markings—researchers regard information from simulators as informative for understanding real driving behavior.

In addition to standard driving behaviors (e.g., speed, staying within a lane, lane position variance), participants' head and eye movements can be monitored in simulations (Campagne, Pebayle, and Muzet 2005) and in real vehicles (Sodhi et al. 2002). Head and eye movements are informative about where attention is being directed. Eye movements are commonly examined in driving situations to determine the effects of distractions on drivers' fixations, or to examine characteristics of adaptive scanning behaviors. For instance, Santos et al. (2005) reported that the effects of secondary tasks on eye movements tend to be similar in real vehicles and in driving simulators. Underwood, Chapman, Bowden, and Crundall (2002) noted that novice drivers tended to show restricted horizontal scan paths, as demonstrated by smaller horizontal scanning variance. These results suggest that novice and more experienced drivers may show qualitatively different scanning behaviors, even within a driving simulator.

Although simulations can be very useful in examining driving behavior, there is at least one negative complication: cybersickness. Some people experience nausea, headaches, and other discomforts when driving in a simulator. The symptoms can be quite severe, especially with some lower fidelity simulators. Stanney and Salvendy (1998) reported that 80% to 95% of participants experience some symptoms of cybersickness after being exposed to virtual reality environments, and 5% to 30% of participants elect to end participation early. Although certain steps can be taken to alleviate symptoms (see Duh, Parker, and Furness 2004), the most effective strategy is to gradually initiate participants to the simulator by initially limiting their exposure time (LaViola 2000) and scheduling several breaks early in training.

These findings suggest the following:

- Simulators that provide wide fields of view and haptic feedback tend to offer a higher sense of realism.
- Participants behave similarly in driving simulators as they do in real vehicles.
- Examining eye movements is useful to determine where in the scene operators are directing their attention and what effects distractors have on people's normal scanning patterns. Some research has found differences in scanning behavior as a function of driving experience.

GENERAL METHODOLOGY

Participants

Iowa State University researchers randomly selected 200 full-time Iowa DOT snowplow operators from a list of 1098 current maintenance workers, with a representative number chosen from the six districts in Iowa. Half of the operators were randomly selected to be in the control group, and the other half comprised the experimental group. These groups were matched in terms of age, district, number of accidents in the past five years and approximate amount of snowplowing and truck driving experience. Operators were informed that they would be receiving mandatory training, but that they had the option of participating in the study to assess training effectiveness. Overall, 174 operators, 84 in the experimental group and 90 in the control group, elected to participate. Scheduling conflicts on the part the researchers or Iowa DOT trainers prevented 24 of these individuals from participating. Only two operators elected not to begin or complete the experiment once they arrived at the facilities. There were 2 females in the sample. The estimated mean age of the sample was 47.6 years (actual age was not provided, rather the age was reported in 10-year increments) and the average years of experience as a snowplow operator was 12.6 (SD = 7.41).

The Simulator

Participants received training in the TranSim VS III truck and snowplow simulator. The VS III has a 180-degree horizontal and 37-degree vertical viewing area at 34 inches from the center screen, and it is comprised of three 1024 x 768 monitors with a refresh rate of 70 Hz. The driving apparatus is similar to what would be found in a typical truck: functional brake, clutch and accelerator pedals, radio, transmission, and digit gauges, all of which can be tailored to the requirements of the trainer. Although a clutch pedal was present, all simulations used an automatic transmission. The simulator also outputs digital sound designed to mimic normal operating sounds, such as engine and exterior noises. Tactile transducers under the drivers' seat provided simulated road vibration. Finally, two graphical rear view mirrors were displayed on the bottom corners of the central screen (spot mirrors), and two graphical adjustable side mirrors were displayed on the outer half of the left and right screens (wing mirrors). A picture of the interior of the simulator is show in Figure 1.

Figure 1. Photo of the simulator

Three weeks into the study, certain parts of the driving simulator were upgraded: more snow was added to the edges of the windshield, and a few additional cars were added to the test scenario. Also as part of the upgrade, some of the tests for the performance measures were changed or excluded. Unless otherwise stated, the performance statistics reported were only included from the 136 participants who completed the training after the upgrade.

In order to track the participants' eyes and head movements, cameras were mounted on the very top of the simulator, approximately six inches above the central monitor, so as not to interfere with participants' performance in the simulator. The tracking was accomplished by running FaceLab (version 3.2) software. Only head and eye movements for the experimental scenarios were recorded.

Procedure

Training was conducted primarily at the main District One facility in Ames, Iowa. Typically, participants were trained in groups of two in a session that lasted for approximately four hours (8 a.m.–noon or 12:30 p.m.–4:30 p.m.). Upon arriving at the facility, the participants were introduced to the trainer, an experienced Iowa DOT snowplow operator, and shown the simulator. The trainer then drove a 3-minute scenario down a sparsely populated rural road to acquaint the participants with the visuals and with the auditory feedback. Next, participants were asked to sign an informed consent document if they wished to participate in the experiment. As mentioned above, only one individual declined to participate, and one operator later elected to withdraw from the study. Operators who gave their consent to participate then drove an introductory 3-minute scenario through a snowy highway to become acclimated with the controls and handling of the simulator. No data were collected during this orientation drive.

Immediately after the orientation drive, participants completed their first set of questionnaires. This set included (1) the NEO-Five Factor Inventory (NEO-FFI), which is based on the Five Factor Model developed by Costa and McCrae (1985); (2) a modified version of Zuckerman's (1979) Sensation Seeking scale (form V); and (3) Witmer and Singer's (1998) immersive tendencies questionnaire. The questionnaires are described in more detail in a subsequent section of the manuscript.

While participants were completing these questionnaires, they took turns driving the *first* experimental scenario. The scenario involved merging onto an interstate in snowy conditions and plowing snow for approximately 10 minutes. The simulated vehicle had a right plow and right wing, although participants did not control this equipment. At various times during the scenario, participants had to pass slower moving vehicles in the right lane, while avoiding striking them with the right wing, and then merge back into the right lane to avoid faster moving vehicles approaching from the rear. Subjects were instructed to operate the simulator in the same manner as they would operate a snowplow while removing snow on a real highway.

After completing the first set of questionnaires and the first experimental scenario, operators assigned to the control group immediately completed the second experimental scenario. The second scenario was identical to the first. Following completion of the second drive, the control group began their training. Participants in the experimental group received training before completing their second experimental scenario. This design, in which the only difference between the control and experimental groups was whether their second experimental drive occurred before or after training, allowed us to determine whether the training had any immediate impact on performance, fixation behaviors, or both. Additionally, this methodology allowed all operators to receive simulator training rather than just those in the experimental condition.

The training consisted of three parts: a lecture that included PowerPoint slides, a computer exercise, and a simulator exercise. The trainer began by giving a 20- to 25-minute lecture, which included a PowerPoint presentation. The lecture focused on the importance of being aware of the space around the vehicle and included a discussion of the S.I.P.D.E. (Scan-Identify-Predict-Decide-Execute) method (see Appendix A). After the PowerPoint lecture, each participant completed a 5- to 10-minute driving scenario in the simulator in which he or she was encouraged to employ the techniques learned during the lecture. The trainer would sporadically ask the participant questions to ascertain whether or not he or she was using the information presented during the lecture. Next, participants completed a self-paced computer exercise in which they watched short video clips that contained information about passing vehicles, speed management, and space management. After each series of clips, participants answered multiple choice questions and received feedback regarding their answers. Finally, participants completed an additional 5- to 10-minute driving scenario in which they were again instructed to employ the techniques that they learned during the PowerPoint lecture and the computer exercise. Training concluded with the trainer giving a 5- to 10-minute summary of the information covered during training.

Once training ended, participants in the control group completed their second set of questionnaires. Participants in the experimental group drove the second experimental scenario before completing the questionnaires. The questionnaires included a modified version of Witmer and Singer's (1998) presence questionnaire, Kennedy et al.'s (1993) simulator sickness questionnaire (SSQ), and a modified version of the questionnaire used by Strayer et al. (2004) for the Utah DOT snowplow study. The questionnaires and modifications are described in a subsequent section of the manuscript.

As part of training, the following was achieved:

- Operators received simulator, computer-based, and lecture training. Thus, people who learn best from different forms of instruction should all benefit from training.
- Participants completed various personality, realism, and training questionnaires to ascertain whether a) operators showed any personality differences from the general population, b) more experienced operators showed any personality differences from less experienced operators, and c) operators deemed the driving simulator to be representative of a real snowplow.
- By dividing operators into two groups, by matching them on various demographic characteristics, and by having only one group complete training before their second drive, we could also determine what effects training had on participants' performance in the simulator.

RESULTS AND DISCUSSION

Characteristics of Iowa DOT Snowplow Operators

One of the purposes of the study was to determine the characteristics of Iowa DOT snowplow operators. Because personality variables are sometimes good predictors of success in or satisfaction with one's occupation (Barrick and Mount 1991; Tett, Jackson, and Rothestein 1991), three separate personality inventories were administered prior to training (the NEO-FFI, the sensation seeking scale, and the immersion scale). One hundred seventy participants completed all three of these questionnaires. In the case of missing data (less than 1% of the data), a participant's mean score for the scale or subscale was substituted for the missing value. For a more in depth discussion on the constructs measured in each questionnaire, see Masciocchi, Dark, and Parkhurst (2006).

Demographics

Just over 1% (2 out of 174) of the participants were women. Figure 2 shows the number of operators from each age category who participated in the study compared to the population of Iowa DOT snowplow operators. As would be expected given the selection procedure, the selected operators were representative of the population of snowplow operators with respect to age.

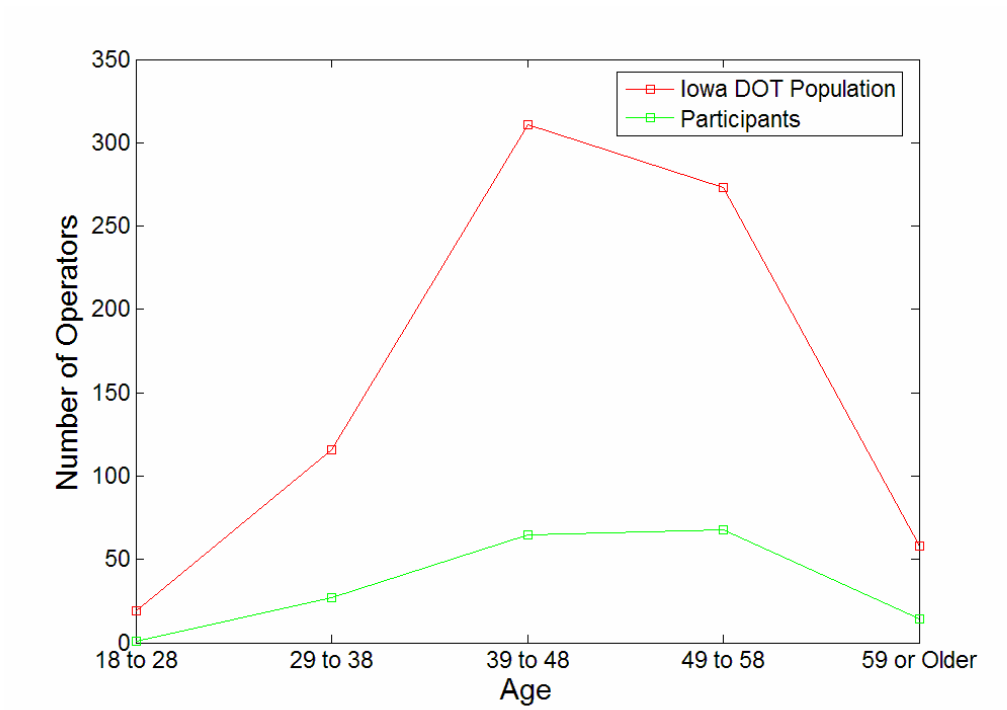


Figure 2. Age of Iowa DOT operators and selected participants

Years of snowplow experience ranged from 0 to 30 years. For purposes of analysis, participants were divided into three categories based on their years of snowplow experience:

- “Low-experience” operators were defined as those with between 0 and 5 years of snowplowing experience (43 operators)
- “Medium-experience” operators were defined as those with between 6 and 15 years of experience (62 operators)
- “High-experience” operators were defined as those with 16 or more years of experience (69 operators)

Figure 3 shows the number of operators in each of the four age groups for the three levels of snowplowing experience. Experience categorization had no effect on the training that the operator received. Of interest was whether various dependent measures would vary as a function of experience.

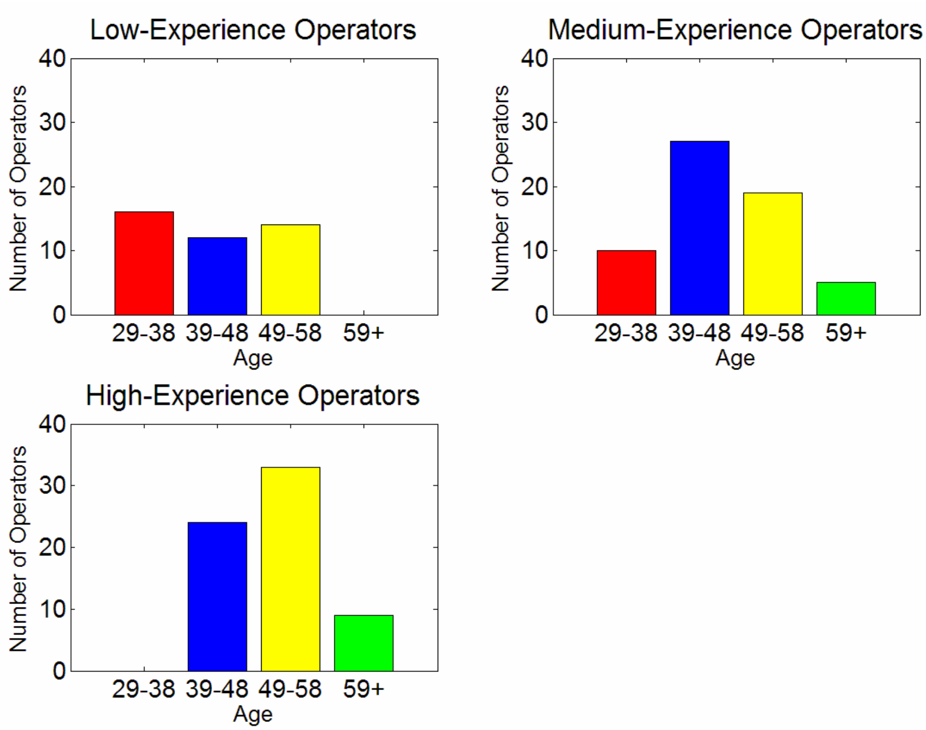


Figure 3. Age breakdown by the three levels of experience

As would be expected, experience was correlated with age category ($r = .39$, $p < .001$). Because personality measures often differ as a function of age, whenever a reliable effect of experience was found in the Analysis of Variance (ANOVA), an Analysis of Covariance (ANCOVA) was run with age category included as a covariate in order to determine whether the effect of experience remained when age was statistically controlled in this way. We chose this method rather than ANCOVA to begin with because of the ordinal nature of the age measure in terms of decades instead of years. We were unsure of the extent to which a decades-based measure might lower the statistical power of our analyses.

NEO-Five Factor Inventory

The first personality questionnaire was the NEO-Five Factor Inventory (NEO-FFI), which is based on the Five Factor Model developed by Costa and McCrae (1985). It is designed to measure personality dispositions in five subscales: neuroticism, extraversion, openness, agreeableness and conscientiousness. The NEO-FFI contains 60 statements (12 for each sub scale). Participants used a five-point scale, ranging from strongly disagree to strongly agree, to indicate the extent to which each statement is a good characterization of their tendencies. The actual questionnaire is in Appendix B. Sample questions are listed below:

- I like to have a lot of people around me (extraversion)
- I often try new and foreign foods (openness)
- I work hard to accomplish my goals (conscientiousness)
- I try to be courteous to everyone I meet (agreeableness)
- I often feel inferior to others (neuroticism)

The first analysis examined whether personality profiles differed as a result of experience level. Of particular interest was whether high-experience participants might be different from low-experience participants. We reasoned that any such difference might be useful in determining the type of person likely to remain a snowplow operator for a longer period of time. The second analysis examined whether any particular personality scale or set of scales characterized snowplow operators in comparison to the general population. We reasoned that any such difference might be useful in selecting persons to be hired as snowplow operators.

Figure 4 shows the mean scores on each of the five sub scales for low-experience, medium-experience, and high-experience participants. Separate one-way ANOVAs were run on scores in each of the five domains in order to determine whether there were differences as a function of experience. Only the openness domain showed a reliable difference ($F(2,167) = 3.43$, MSE = 29.3, $p < .05$). Least Significant Differences (LSD) comparisons among the experience groups showed that the low-experience participants were more open to new experiences than the medium- and high-experience participants, who did not differ. However, when age category was included as a covariate in an ANCOVA, it was a reliable covariate and the effect of experience was no longer apparent. This suggests that the observed difference was reflecting age more than experience level. Thus, the personality subscales measured by the NEO-FFI did not differentiate among individuals in the sample in terms of experience.

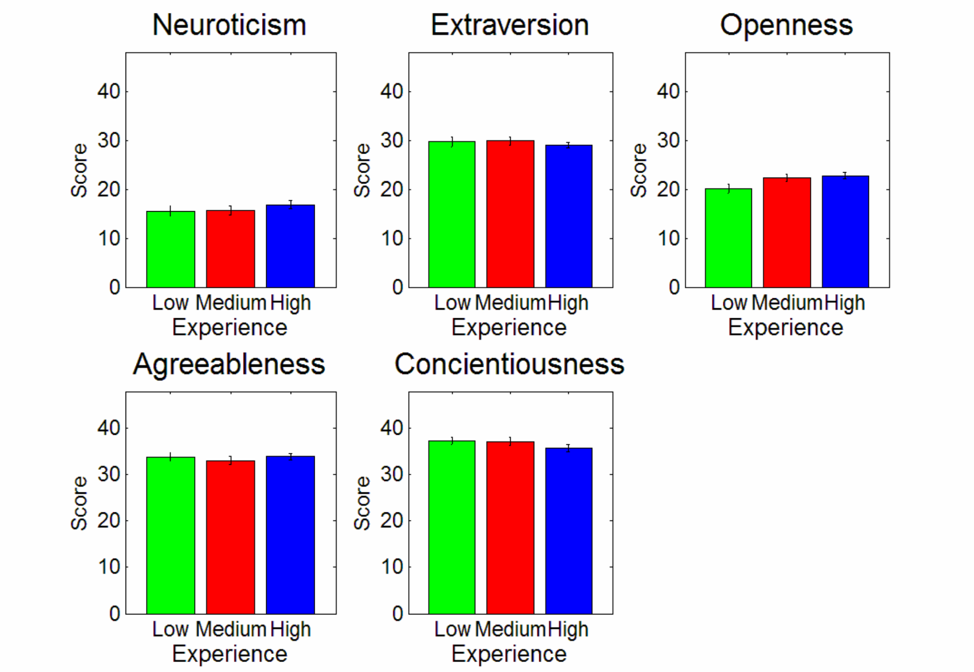


Figure 4. Mean NEO-FFI score on each subscale for the three experience groups

The NEO-FFI scores of the snowplow operators measured in our sample were then compared to those of a normative sample of 500 American males as reported in Rolland, Parker, and Stumpf (1998). Table 1 shows the normative scores and standard deviations reported on the five NEO-FFI domains along with the scores of the Iowa DOT participants collapsed over experience category. Except for the openness trait, the observed differences generally were rather small, but they were reliable. The comparisons treated the normative value as a population mean against which the participants' mean was compared. Snowplow operators were less neurotic, more extraverted, less open, more agreeable, and more conscientious (all *t*s > 2.73, all *p*s < .05). Although the data suggest that some NEO-FFI traits might provide information that could be used in the selection of snowplow operators, it is highly likely that the snowplow operators differed from the normative sample in many ways, particularly in terms of demographics. A more appropriate control group would be a sample of Iowa DOT employees who are not snowplow operators. These individuals would be more similar to the sample of snowplow operators on several important factors, particularly age, race, and income. This more appropriate matched sample would be needed before the usefulness of the NEO-FFI as a selection criterion could be determined.

Table 1. Normative American male scores and operator scores on each NEO-FFI subscale

Authors	Neuroticism	Extraversion	Openness	Agreeableness	Conscientiousness
Rolland et al. (1998)	17.6 (7.46)	27.2 (5.85)	27.1 (5.82)	31.8 (5.03)	34.1 (5.95)
Present Study	16.1 (7.16)	29.5 (6.06)	22.0 (5.49)	33.6 (6.15)	36.7 (6.08)

Sensation Seeking

Sensation-seeking tendencies were explored by Zuckerman (1979), who described the personality construct of sensation seeking as a tendency to seek out novel or exciting experiences for the sake of enjoying the activity. The sensation-seeking questionnaire has four subscales: thrill and adventure seeking, experience seeking, disinhibition, and boredom susceptibility. We used a modified version of Form V of the questionnaire. Form V contains 40 questions, 10 questions for each subscale. However, seven questions were removed from the disinhibition subscale and three were removed from the experience seeking subscale because they asked about drug or alcohol use or other controversial behaviors. The modified questionnaire is shown in Appendix C. For each question, the participants choose which of two statements is a better description of their behaviors or preferences. Example questions are listed below:

- Thrill Seeking
  - A sensible person avoids activities that are dangerous
  - I sometimes like to do things that are a little frightening
- Boredom Susceptibility
  - I get bored seeing the same old faces
  - I like the comfortable familiarity of everyday faces
- Disinhibition
  - I like uninhibited parties
  - I prefer quiet parties with good conversation
- Experience Seeking
  - I prefer "down to earth" kinds of people as friends
  - I would like to make unusual friends

The first analysis examined whether more experienced participants have different sensation seeking profiles than less experienced operators. As with the NEO-FFI, we reasoned that any such difference might be useful in determining the type of person likely to remain a snowplow operator for a longer period of time. The second analysis compared snowplow operators to the general population.

The mean scores for the two complete subscales—thrill and adventure seeking and boredom susceptibility—are shown in Figures 5 and 6. A one-way ANOVA on each of the subscale scores with experience as a between-subjects variable showed a difference on the thrill and adventure seeking subscale ($F(2, 167) = 3.33$, MSE = 8.60, $p < .05$), in which high-experience participants reported less thrill- and adventure-seeking tendencies than the middle-experience participants (LSD Comparisons). Zuckerman (1979) reported that older individuals in general show lower sensation-seeking scores than younger individuals. When age category was included as a covariate, it did not produce a reliable difference, but experience effect was no longer reliable ($p > .11$).

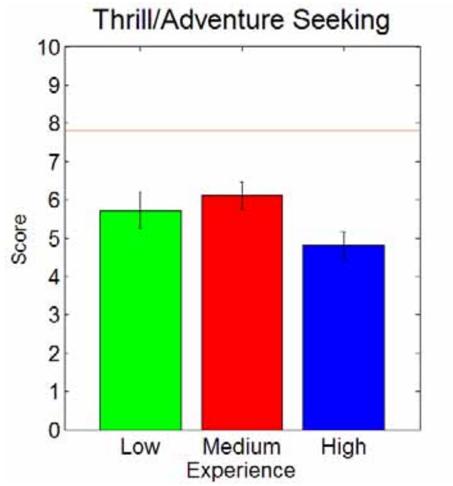


Figure 5. Mean thrill and adventure seeking scores for the three experience groups

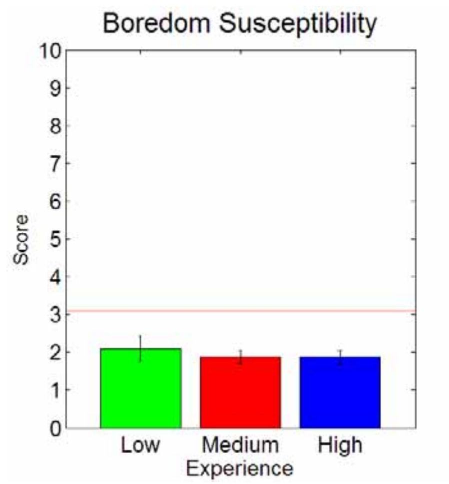


Figure 6. Mean boredom susceptibility scores for the three experience groups

There was also some evidence that Iowa DOT snowplow operators show less sensation seeking tendencies overall than the general population. The red bars in Figures 4 and 5 depict the mean scores for a group of 377 American males on the thrill and adventure seeking and boredom susceptibility subscales, respectively (Zuckerman 1979); these were the two subscales that were presented in their entirety in the current study. For both traits, all groups of participants, regardless of experience level, reported less sensation seeking than the normative average, all ts > 4.70 and all ps $< .01$ for thrill and adventure seeking and all ts > 3.00 and all ps $< .05$ for boredom susceptibility. The comparisons treated the normative value as a population mean against which the participants' mean was compared. As was indicated in the discussion of the NEO-FFI traits, the current data are only suggestive. An appropriate control group is needed before the usefulness of sensation seeking measures as a selection tool can be determined.

Given the correlation between age and years of experience and the fact that the Iowa DOT operators were older in general than the normative sample, the interpretation of this difference must be made with caution. Iowa DOT operators may have lower sensation tendencies than the

American males in other research samples, but this difference may also be a function of age. It is likely that the Iowa DOT sample was older than the individuals examined by Zuckerman (1979).

Immersion

Immersion is commonly defined as the ability to become enveloped or involved in an environment or task (Witmer and Singer 1998). It is a general characteristic not directly tied to virtual reality environments. High immersion scores also imply the ability to block out or ignore task-irrelevant distractions. We examined whether immersion would vary as a function of experience since, we reasoned, individuals with lower immersive tendencies may not become as involved or focused in training; therefore, they might not experience as much of a benefit from it.

The immersion questionnaire (Witmer and Singer 1998) contained 29 questions, a majority of which are related to one of three domains: focus, involvement, and games. For the majority of questions, participants respond on a scale from 1 (not often/much) to 7 (very often/much) on the extent that the statement typifies their behavior. The questionnaire is shown in Appendix D. Sample questions include the following:

- Do you ever become deeply involved in movies or TV dramas? (involvement)
- How good are you at blocking out external distractions when you are involved in something? (focus)
- How often do you play arcade or video games? (games)

The mean immersion scores for each experience group are shown in Figure 7. One-way ANOVAs with experience as a between-subjects variable showed no differences, $F(2, 167) < 1$. Although one must be cautious when interpreting the lack of a difference, the means suggest that older, high-experienced operators are as likely to become immersed in the simulator as low-experienced operators. To the extent that immersion is a necessary prerequisite to successful simulator training, then, experienced operators should benefit as much from training as less experienced operators.

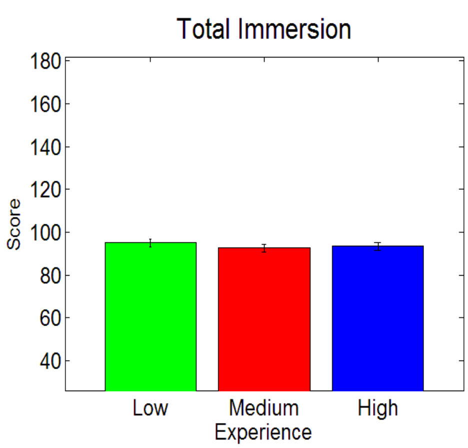


Figure 7. Mean total immersion scores for the three experience groups

Summary of Demographic and Personality Variables

Although they varied in terms of years of experience as a snowplow operator, the participants of this study were comparable on three well-researched personality measures when age differences were taken into account: the NEO-FFI, the sensation seeking scale, and a measure of immersion. While none of the measures currently provides a useful selection tool to determine who might make a good, or at least a long-term, snowplow operator, the immersion measure does suggest that snowplow operators at all experience levels will benefit from simulation training.

Overall, the personality and demographic data show the following:

- The sample of participants selected for this study is similar to the population of Iowa DOT snowplow operators with respect to age.
- There were no differences as a function of experience once age differences were taken into account on measures of personality, sensation seeking, and immersion.
- While there was some indication that Iowa DOT snowplow operators may have some personality differences compared to the general population, a more appropriate control group would be needed to determine whether these differences are informative.
- Based on the fact that operators of all experience levels had similar immersion scores, they should all be able to benefit from simulator training.

Responses to Simulator Training

A major purpose of the study was to determine the effectiveness of simulator training for snowplow operators. In order for the simulator to be a useful tool in training novices and in providing experienced operators with a way to practice their skills throughout the year, it must provide a realistic driving environment. The judged realism of the simulator could vary as a function of experience (e.g., high-experience operators may find the simulator to be more or less realistic than low-experience operators), making the simulation more or less useful for operators at a particular experience level. We examined participants' judgments about simulator realism through their responses to the presence questionnaire, the simulator sickness questionnaire, and direct questions about simulator realism. We also examined participants' satisfaction with the entire training experience.

Presence

Presence is directly related to immersion in a virtual reality environment. Witmer and Singer (1998) defined presence as "the subjective experience of being in one place or environment, even when one is physically situated in another." For example, although operators were physically located in the interior of a trailer during training, those with high levels of presence may have had the experience of actually driving on a real highway. We examined whether experience had any effect on level of presence, with the implication that low levels of presence may result in less satisfaction with the driving simulator or simulator training. We thought it is possible that more experienced operators could be more cognizant of inconsistencies between real and simulated snowplow driving, and therefore might experience less presence.

The presence questionnaire (Witmer and Singer 1998) contains 32 questions from 4 domains: control, sensory, distraction, and realism. Participants respond on a scale from 1 (not often/much) to 7 (very often/much) on the extent that each statement typifies their experience in the simulator. The wording of some items was modified slightly to make them more applicable to the current study involving snowplows. The modified presence questionnaire is shown in Appendix E. Sample questions are listed below:

- How natural did the driving simulator seem (control)?
- How much did the scenery of the simulator make you feel like you were really driving a snowplow (sensory)?
- How natural were the driving controls (distraction)?
- How similar was driving the simulator to driving a real snowplow (realism/control)?

The mean presence scores for each experience group are shown in Figure 8. A one-way ANOVA with experience as a between-subjects variable showed no differences between levels of experience. In other words, operators of all experience levels reported experiencing a similar amount of presence. Moreover, a one-sample t-test comparing participants' average response to a score of neutral (i.e., a score of 4 for each question) revealed that scores were significantly greater than this value ($t(169) = 8.14$, $p<.001$). Thus, participants reported an above average sense of presence.

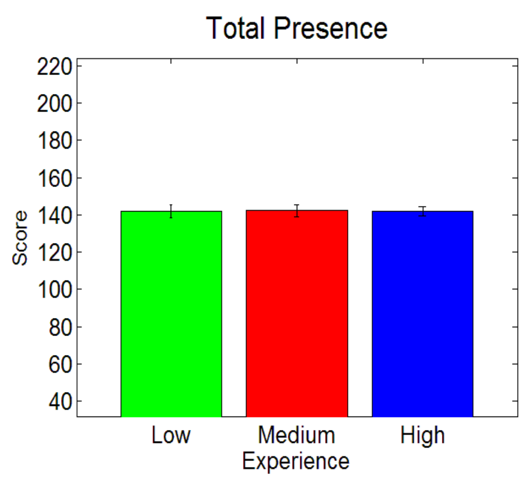


Figure 8. Mean total presence scores for the three experience groups

Simulator Sickness

Simulator sickness is always a concern when virtual reality devices are used for training because people who experience simulator sickness may not be able to concentrate as fully on training or they may be unable to finish training. Of particular concern is that older individuals may be at increased risk for simulator sickness (Arms and Cerney 2005). The simulator sickness questionnaire (SSQ), which was developed by Kennedy et al. (1993), was used to measure participants' reported level of simulator sickness after completing training. The SSQ contains 16 questions regarding potential symptoms of simulator sickness. Participants are instructed to report, on a 0 (none) to 3 (severe) scale, the extent to which they experience each of an assortment of symptoms. The reported values are used to compute a total SSQ severity score that ranges from 0 to 180. The full questionnaire and scoring procedures for calculating the total weighted simulator sickness score are shown in Appendix F. Sample symptoms include the following:

- Headache
- Nausea
- Blurred vision

The mean SSQ scores for subjects from each experience group are shown in Figure 9. A one-way ANOVA with level of experience as a between-subjects variable showed no effect of level of experience ($F(2, 167) = 1.62$, MSE = 1099.3, $p>.20$). However, an independent samples t-test between operators with high versus low levels of experience was marginally significant ($t(107) = 1.80$, $p=.075$), suggesting that more experienced drivers did show a tendency to experience more simulator sickness. The most reasonable explanation for this finding is that older drivers experienced more simulator sickness than younger drivers. Thus, there is some indication that simulator sickness was related to age, as has been shown by others (e.g., Arms and Cerney 2005).

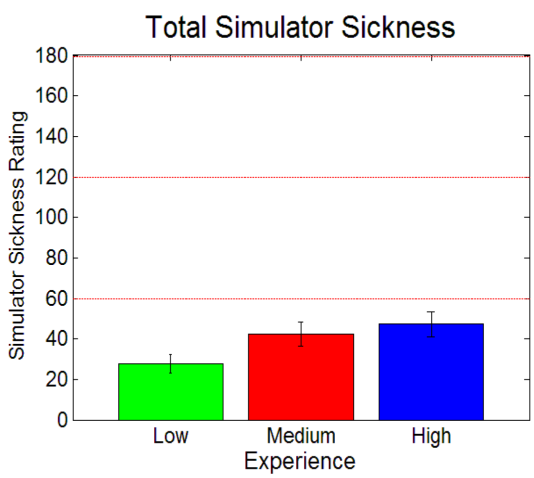


Figure 9. Mean weighted total simulator sickness scores for the three experience groups

The most important finding, however, is that participants' simulator sickness ratings were relatively low. A weighted score of approximately 60 corresponds to participants reporting slight discomfort for all questions. (A weighted score of approximately 120 corresponds to moderate discomfort for all symptoms, and a weighted score of approximately 180 corresponds to severe discomfort for all symptoms.) The means of all experience groups were well below this rank (all *t*s > 5.86, all *p*s < .001). (The comparisons treated a score of 60 as a population mean against which the participants' mean was compared.) Additionally, only 5 out of a total of 174 participants were unable to complete the simulator training due to excessive simulator sickness. Thus, it does not appear that simulator sickness is an obstacle to the use of this simulator in training Iowa DOT operators.

Simulator Realism

Questions regarding the realism of the TransSim VS III were added to the end of the training questionnaire used by Strayer et al. (2004) along with free response questions regarding training satisfaction. The questions addressed the realism of certain aspects of the driving simulator: the scenery (Question 1), the visuals of other vehicles (Question 2), the location (Question 3) and functionality (Question 4) of the mirrors, the behavior of other vehicles (Question 5), the location (Question 6) and responsiveness (Question 7) of the simulator's controls, and whether the operator thought driving the simulator approximated the experience of driving a real snowplow (Question 8). The questions can be found in Appendix G. Realism was rated on a scale of 1 (strongly disagree) to 7 (strongly agree).

The mean responses to each of the eight aspects of the simulator for all operators are shown in Figure 10. One-way ANOVAs on the responses with experience as a between-subjects variable showed no differences as a function of experience. Next, the mean response to each question was statistically compared to 4, the neutral point, to determine whether the judged aspect was considered to be realistically or unrealistically presented in the simulator. All aspects of the simulator, except for the overall realism rating were judged to be realistic (all *t*s >3.80, all *p*s < .01). The real snowplow approximation rating was neither above nor below the neutral point. Thus, the responses to the direct question regarding realism show that participants found the visual qualities, location, and functionality of the equipment to be similar to those in a real driving situation.

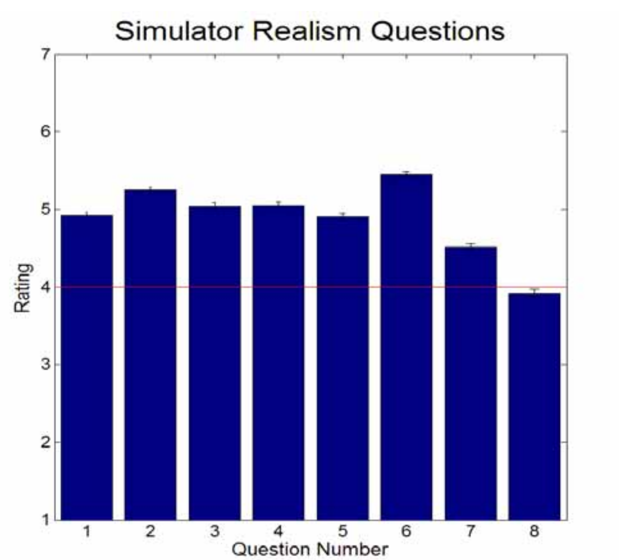


Figure 10. Mean simulator realism ratings for the eight realism questions

Training

The major part of the final questionnaire was based on the questionnaire used by Strayers et al. (2004) in their simulator fidelity study of Utah DOT operators. The complete questionnaire is shown in Appendix G. Subjects responded on a scale of 1 (Strongly Disagree) to 7 (Strongly

Agree) to statements about five aspects of training: simulator training (5 statements), computer training (5 statements), lectures (4 statements), the trainer (4 statements), and the value of training (4 statements). The mean response to each of the five categories for each experience group is shown in Figure 11. A one-way ANOVA on each category with experience as a between-subjects variable showed no differences due to experience. Thus, more experienced operators tended to rate training as highly as less experienced operators. This is an important finding, as it suggests that all operators regardless of level of experience felt that they benefited from, or at least enjoyed, training.

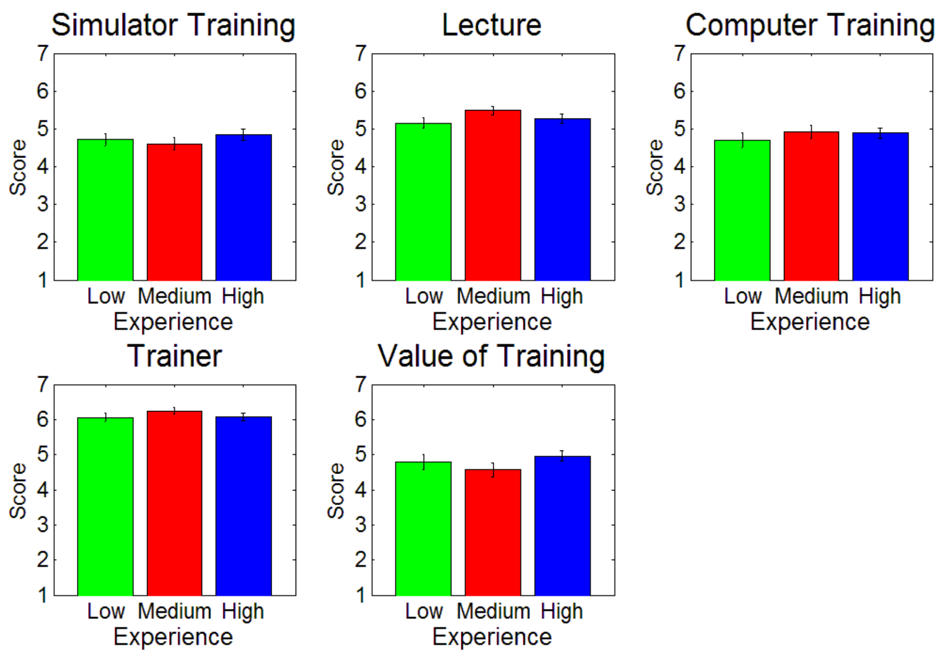


Figure 11. Mean ratings for the five training categories for the three experience groups

The mean response to each category was statistically compared to 4, the neutral point, to determine whether participants' judgments were positive. All categories received liking ratings (all *t*s >7.18, all *p*s < .001); participants liked all aspects of training, especially those that involved the trainer (lecture and trainer). The uniformity of positive responses is conveyed in Figure 12, which presents the average proportion of participants who responded with dislike scores (1 to 3), neutral scores (4), or like scores (5 to 7) to each item, comprising the five aspects of training. In each case, over 55% of respondents liked that aspect of training or found it valuable.

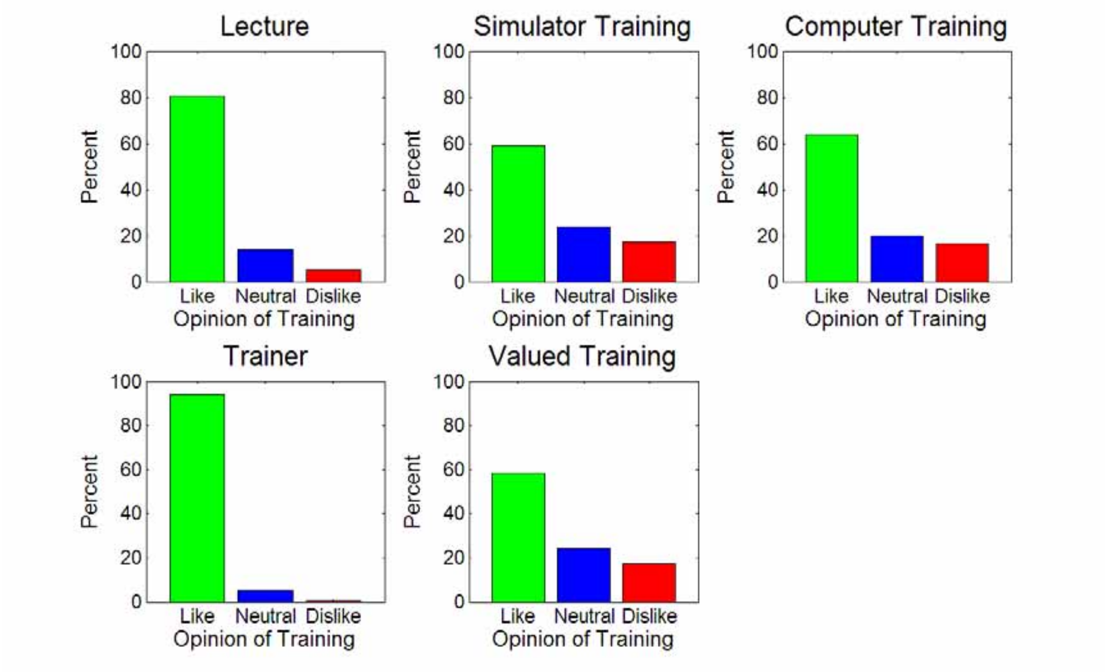


Figure 12. Percent of participants who liked, were neutral, or disliked the five aspects of training

Qualitative Training Questions

The six qualitative questions and a selection of the more common responses are shown in Table 2. The open-ended responses corroborated several aspects of the quantitative data. As implied by the lower score on the "driving a real snowplow" realism question, many operators commented that the driving simulator lacked some of the equipment that is normally present on snowplows, such as plow and wing controls. Others commented that the calibration of engine sound and speed seemed off, and sounds such as the scraping of the plow on the roadway, that are typically present during snowplowing, were absent in the simulator. Finally, in spite of the apparent discrepancies between the simulator and actual snowplow driving, many operators stated that they thought the driving simulator was the most useful part of training. Thus, although not every operator enjoyed the simulator training, many rated it as the most enjoyable or useful part of training.

Table 2. Selected responses for the qualitative questions portion of training questionnaire

Question	Selected Operators' Responses
1) Is there anything that can be done to make the scenery better?	- Difficult to judge distance of oncoming traffic - More snow and slush on windshield during driving - Majority thought it was fine or had no comments
2) Is there anything that can be done to make the sound better?	- Better calibration between engine sound and speed - Add sounds from plow equipment (e.g., scraping on highway) - Louder engine noise
3) Is there anything that can be done to make training better?	- Add controls for plow and wing - More time driving the simulator - Night time snowplowing
4) What was the most useful part of training?	- All: 15% - Simulator/Driving: 40% - Lecture/SIPDE/PowerPoint: 30% - Computer: 10% - Miscellaneous: 5%
5) What was the least useful part of training?	- Simulator: 25% - Computer: 20% - Lecture: 5% - None (It was all useful): 35% - Miscellaneous: 15%
6) Do you have any additional comments?	- Add plow and wing controls - Simulator accelerates and brakes too fast - Steering is too sensitive - Need more feedback on driving performance - Training was useful, especially for new employees

Summary of Simulator Training and Realism Results

Similar to the immersion scores, operators from all experience levels reported similar amounts of presence in the simulator, suggesting that they should all benefit from simulator training. Operators reported in the quantitative and qualitative sections that they thought the simulator was realistic, although they made some suggestions on aspects that could be improved. Responses to training were also quite positive, and participants reported that they particularly enjoyed the simulator and lecture portions. Finally, simulator sickness scores were overall quite low. Mean simulator sickness scores fell below the level of slight discomfort, and less than 3% of participants in this study elected to end training early due to excessive cybersickness.

Operators' responses to simulator realism and training questionnaires revealed the following:

- Presence scores were consistent for participants across levels of experience, suggesting all operators, regardless of amount of snowplowing experience, should be able to benefit from simulator training.
- Training was well received by operators from all levels of experience, and realism scores were typically also high.
- Importantly, simulator sickness scores were relatively low, suggesting that this simulator is appropriate for training a large population of individuals.

Performance Analysis

In this section, we examined whether any improvements in simulator performance could be found as a result of participants experiencing this training regimen. It is important to keep in mind that training only tended to last between 1.5 and 2 hours, and during this time, participants drove the simulator for approximately 10 to 20 minutes. Subsequent simulator training provided by Iowa DOT, when additional time does not need to be allocated to administering questionnaires or conducing untrained experimental drives, would almost certainly involve more training time, particularly for driving the simulator. Any differences found as a function of this abridged training regimen, therefore, would be notable. Overall, performance data for participants' first and second drives in the simulator, which were recorded by the driving simulator software, were successfully collected from 124 participants. This number reflects those operators who participated after the simulator upgrade. Three measures of performance were examined: collision rate, average speed, and fuel efficiency.

Critical Failures—Collisions

Critical failures in this study refer to the number of collisions that participants were involved in while driving the simulator. In the tests reported by the simulator software, critical failures also included failure to wear a seatbelt. However, these violations were removed from this examination; thus, critical failures reported in this analysis were limited to accidents. Also, this collisions analysis was conducted on 154 participants, including 30 operators who participated prior to the simulator upgrade. Presumably, minor changes in the visual quality of the simulator should not have lead to any differences on such a broad measure of performance as number of collisions; thus, including these operators was deemed to be appropriate.

A 3 (experience level: low, medium, high) $\times$ 2 (group: experimental, control) $\times$ 2 (time: pretraining drive, postraining drive) ANOVA was conducted on the collision data. Experience level and time were within-subjects variables and group was a between-subjects variable. (This type of analysis will be repeated for subsequent performance, as well as fixation, measurements.) The analysis showed a main effect of group ($F(1, 148) = 4.50$, MSE = .180, $p < .05$) in that participants in the experimental group made more collisions than those in the control group. However, as can be seen in Figure 13, the difference between the groups was apparent on the pretraining drive. The fact that the group effect did not interact with time means that the difference cannot be attributed to training. For an unknown reason, the groups differed prior to the experimental manipulation and that difference remained after the manipulation, since the group effect did not interact with time.

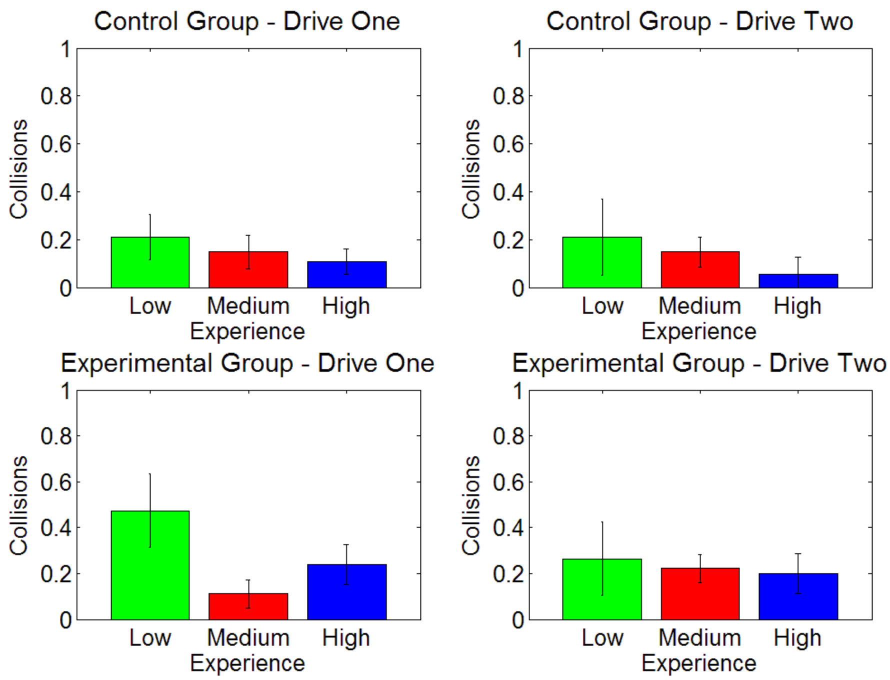


Figure 13. Mean number of collisions for the three experience groups during the first and second drives

The ANOVA also showed a trend towards a main effect of experience level ($F(2, 148) = 2.93$, MSE = .180, $p = .057$). LSD comparisons among the experience groups showed that the low-experience participants were involved in more accidents than the medium- or high-experience groups, who did not differ from each other. Thus, experienced drivers, who presumably had a longer time to refine their skills in the field prior to receiving simulator training, made fewer collisions in the driving simulator than less experienced operators who had not had as much snowplowing experience.

We had expected that training might be more beneficial for less experienced drivers than for more experienced drivers, but the experience level X group X time interaction was not reliable ($F < 1$). Low-experience participants in the experimental group showed more improvement from the first to second drive than the low-experience participants in the control group, but the apparent benefit was not statistically significant. One problematic aspect of the data deserves comment: just what was counted as an accident by the simulator software is unclear. Each of the drives lasted only 10 minutes, but the number of critical accidents ranged from 0 to 3. It may be that variations in the scenario not under participant control contributed to differences in performance. For instance, an “accident” when no other vehicles were programmed to approach within the next few seconds may have yielded a count of one critical error, while an “accident” when another vehicle was programmed to approach in the following few seconds may have yielded a count of more than one critical error (i.e., if it too collided with the operator’s vehicle). The lack of training effects may simply reflect the fact that there was not sufficient change in the two-hour training period to be detected in an immediately following 10 minute drive, but aspects of the simulation programming as just described may also have introduced variability in the data that reduced the power to detect real differences. In support of the latter is the observation that if those participants who have three "accidents" during one of their drives are excluded from the analyses (2 cases), the predicted experience level X group X time interaction is reliable ($F(1, 151) = 4.16$, MSE = .176, $p < .05$), with low-experience drivers showing a reliable benefit of training.

Speed and Fuel Measures

Speed and fuel comparisons were made between operator's first and second experimental drives in the simulator. Figure 14 shows the mean average speed for the three experience levels across their first and second drive, organized by whether they were in the experimental or control group. A pair of 3 (experience) X 2 (group) X 2 (time) ANOVAs were run to determine whether training or experience level had any effect on operators' average speed or average fuel consumption for the first or second drives. First, a main effect of drive was found for average speed ($F(1, 118) = 32.22$, MSE = 10.56, $p < .001$), with operators tending to drive faster on their second drive compared to their first. This was qualified, however, by a reliable interaction with experimental group ($F(1, 118) = 10.70$, MSE = 10.56, $p < .01$). The experimental group drove significantly faster on drive two compared to drive one (mean change = 3.86, SD = 4.77, $t(57) = 6.17$, $p < .001$), while the increase in the control group was smaller (mean change = 0.99, SD = 4.34, $t(65) = 1.84$, $p = .07$).

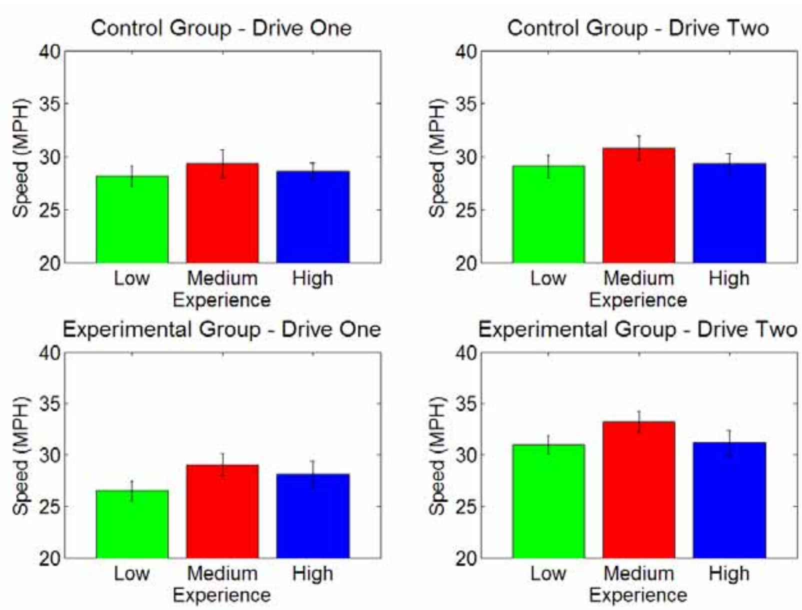


Figure 14. Mean average speed for the three experience groups during the first and second drives

When operators' average fuel consumption data were examined, a statistically significant interaction between experimental condition and drive number was found ($F(1, 118) = 6.83$, MSE = .013, $p = .01$). This showed that participants in the experimental group showed a trend for slightly lower fuel efficiency in drive two compared to drive one (mean change = 0.04, SD = 0.17, $t(57) = 1.75$, $p = .085$), while operators in the control group showed a slight improvement (mean change = 0.045, SD = 0.15, $t(65) = 2.43$, $p < .05$). However, this absolute difference in fuel consumption between the two groups is relatively minor. While it is difficult to determine what a 0.08 average difference between the two groups in fuel consumption as measured by the simulator would translate to under real life estimates, it appears as if operators in the experimental condition tended to drive faster than operators in the control group during their second drive, after they received training, while showing a negligible increase in fuel. Thus, training may have lead to a positive effect on operators' driving performance in the simulator. There was also a main effect of experience ($F(1, 118) = 3.62$, MSE = .050, $p < .05$), with an LSD comparison revealing that low-experience operators had higher fuel efficiency than medium-experience or high-experience operators. The absolute difference was quite small, though, but it may reflect a tendency for low-experience operators to drive more conservatively.

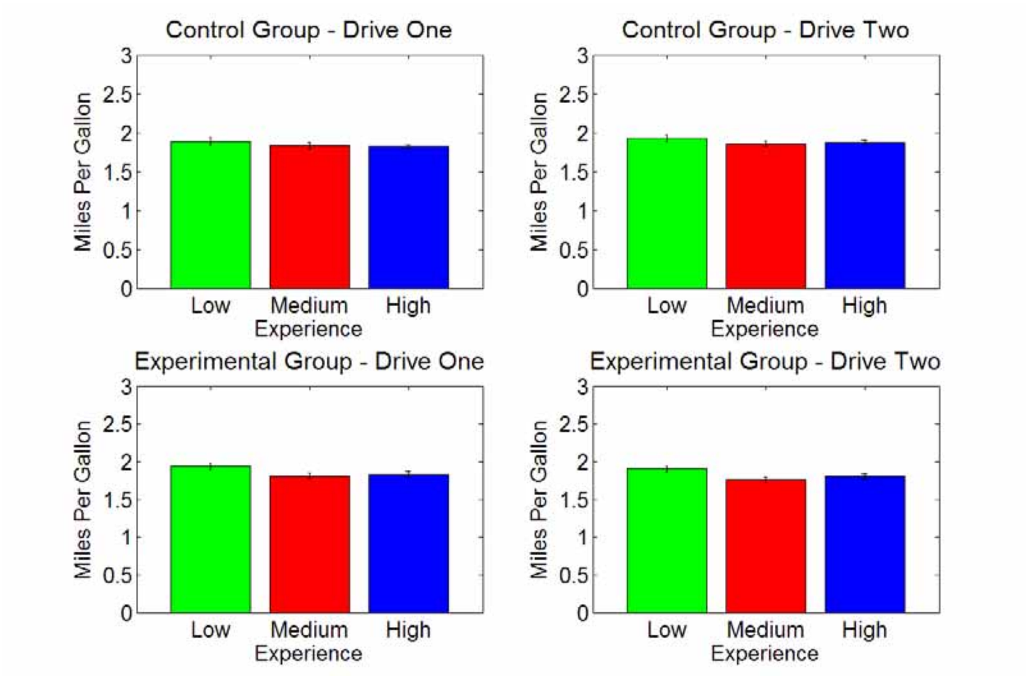


Figure 15. Mean fuel efficiency for the three experience groups during the first and second drives

Summary of Performance Results

There was some indication that operators who received training performed better on their second drive compared to operators who did not receive training. Specifically, these operators drove faster on their second drive while showing a negligible increase in fuel consumption. Additionally, low-experience operators who received training may have shown a tendency to get into fewer collisions than those who did not, but this result is unclear. Finally, low-experience operators were involved in more accidents than medium-experience or high-experience operators.

Taken as a whole, the performance data suggest the following:

- Operators in the experimental group who had training appeared to perform better on their second experimental drive than operators in the control group who had not had training at that point.

Tracking

Participants' fixation behaviors were examined on their first and second experimental drives in the snowplow simulator. Head tracking data were successfully collected for 119 participants, and eye tracking data were collected for 72 participants. (Some of these operators participated prior to the simulator upgrade.) Three categories of tracking analysis were examined: the percentage of fixation time on and number of fixations made to certain critical objects (center screen, right panel, left panel, right spot mirror, left spot mirror, right wing mirror, left wing mirror, and speedometer), horizontal and vertical spread of search, and eye closure.

Object Fixation

The percentage of time that each participant fixated on the center panel was examined as a function of experience. A 3 (experience) X 2 (group) X 2 (time) ANOVA examining the amount of time that operators fixated on the center panel revealed no differences as a function of experimental group, experience, or time (all *p*s > .2). Thus, participants in all groups and levels of experience spent approximately equal amounts of time fixating on the center screen in their first and second drives.

Next, the fixation pattern of operators in each experience group was examined to determine whether training influenced the relative amount of time that operators spent looking at objects other than the center panel. In order to take into account differences among operators in the absolute amount of time spent looking on the center panel, we first determined for each operator the amount of time spent *not* looking at the center panel. We then computed the proportion of this time that was spent on each object. For instance, if an operator spent 20% of the time fixating on objects other than the center panel, the proportion of that amount of time that he or she looked at each object was determined. Multiple 2 (group) X 2 (time) ANOVAs were conducted for each of the critical objects described above. No differences were found for operators in the high-experience group as a function of experimental versus control group or drive number, and some differences were found for operators in the medium-experience group. Specifically, these operators spent less time looking at the right panel ($F(1, 43) = 4.82$, MSE = 353.2, $p < .05$) and tended to spend less time looking at the right wing mirror ($F(1, 43) = 3.19$, MSE = 55.4, $p = .08$). However, a consistent pattern emerged for operators in the low-experience group: these operators spent less time looking at objects on the right side than on the left side of the simulator on their second drive compared to their first. Specifically, low-experience operators spent less time looking at the right panel ($F(1, 30) = 3.08$, MSE = 303.0, $p = .09$) and the right spot mirror ($F(1, 30) = 4.65$, MSE = 102.9, $p < .05$), and they spent more time looking at the left panel ($F(1, 30) = 2.95$, MSE = 77.9, $p = .096$) and the left wing mirror ($F(1, 30) = 4.42$, MSE = 64.6, $p < .05$) (see Figure 16).

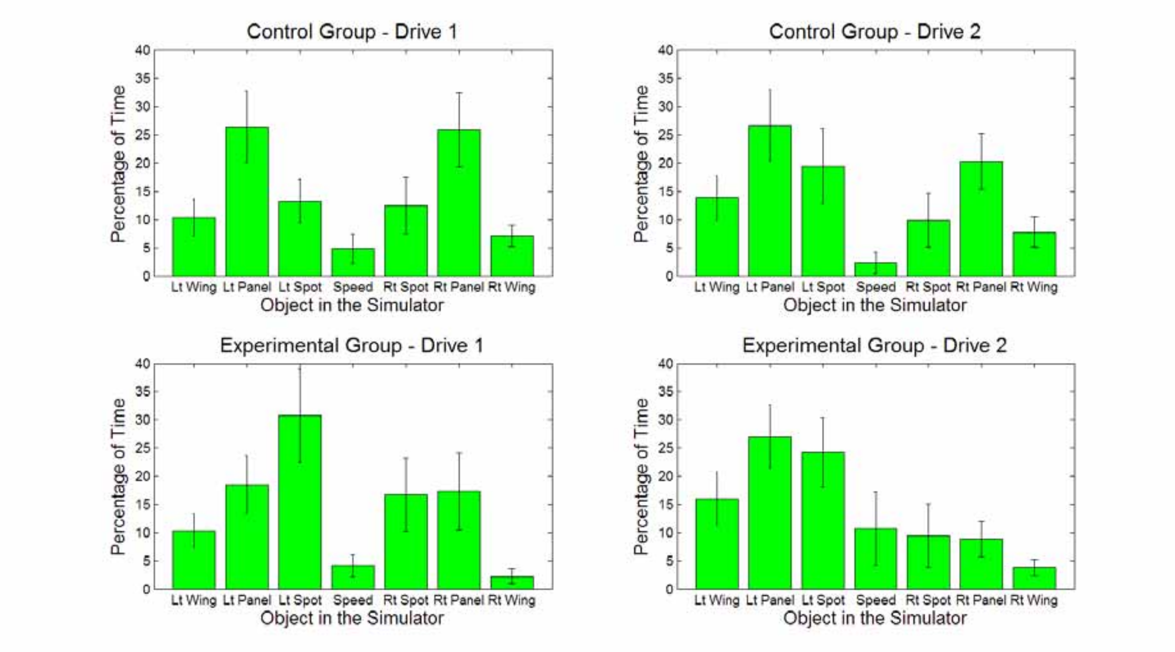


Figure 16. Mean percentage of fixation time on objects outside of the center panel for low-experience drivers in the control and experimental groups on the first and second drives

This pattern was confirmed by examining the percentage of time spent looking outside of the center panel collapsed across objects on the left versus right side (see Figure 17). In other words, we looked at the amount of time operators in the low-experience group spent looking at the left compared to the right side of the center panel on their first and second drives in a 2 (left versus right) X 2 (time) ANOVA. This analysis showed a main effect of location ($F(1, 30) = 5.93$, MSE = 645.4 , $p < .05$), suggesting that operators spent more time looking at the left versus the right side of the snowplow. This was qualified by a reliable location by time interaction ($F(1, 30) = 6.12$, MSE = 645.4, $p < .05$), which showed that operators decreased the amount of time spent looking at the right side on their second drive compared to the first and fixated more often on the left side during their second drive compared to the first.

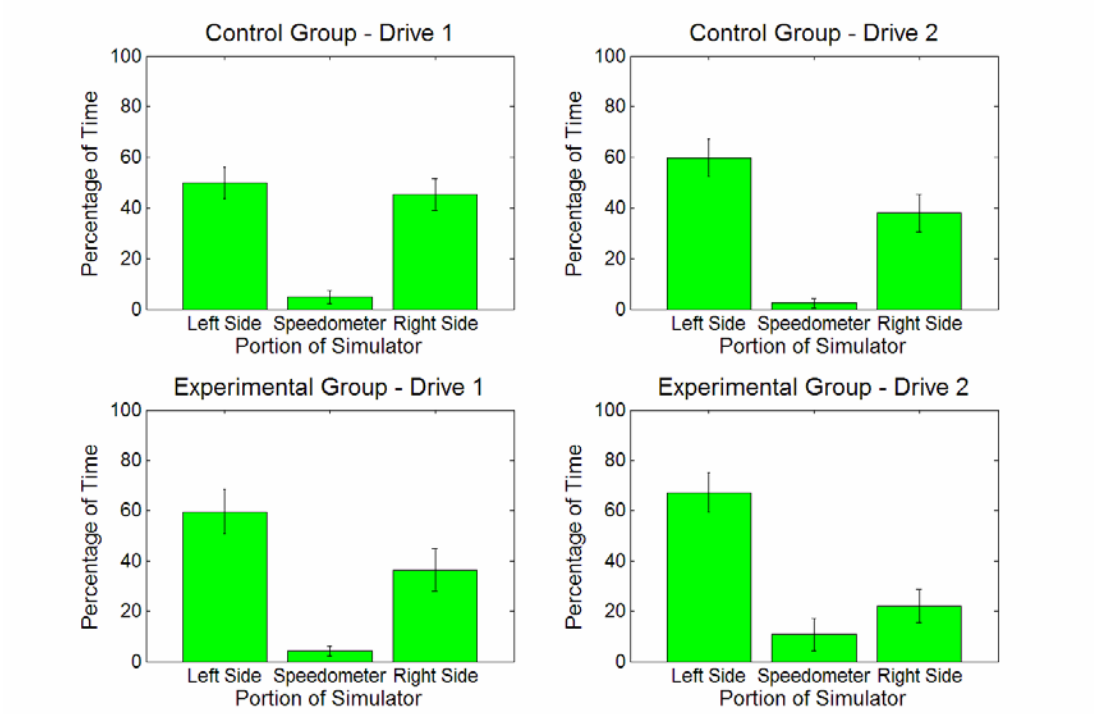


Figure 17. Mean percentage of fixation time on left and right sides for low-experience drivers in the control and experimental groups on the first and second drives

Finally, an analysis of object fixation changes was conducted in an attempt to provide converging evidence for these findings. An object fixation change in this study is defined by a fixation change from one object to another. For example, if a participant looks from the left spot mirror to the center panel, this would count as a fixation change away from the left spot mirror. Multiple 2 (group) X 2 (time) ANOVAs were conducted to examine the number of fixations on each critical object in the simulator (see Figure 18). Overall, the general pattern that was found for time spent fixating on the objects outside of the center panel was replicated in this analysis. Specifically, low-experience operators made fewer fixations to the right spot mirror ($F(1, 31) = 6.15$, MSE = .002, $p < .05$) and more fixations to the left panel ($F(1, 31) = 7.48$, MSE = .003, $p = .01$) and the left wing mirror ($F(1, 31) = 8.48$, MSE = .002, $p < .01$) on their second drive compared to their first drive. A main effect of time was also found for the middle panel ($F(1, 31) = 4.61$, MSE = .003, $p < .05$), demonstrating that operators made fewer fixations to the center panel on their second drive. Additionally, the right spot mirror fixation reduction was qualified by a significant group interaction ($F(1, 31) = 6.51$, MSE = .002, $p < .05$), which showed that a greater reduction in number of fixations occurred for operators in the experimental group than in the control group. However, since the probability of obtaining reliable differences increases with the number of tests conducted, this may not reflect a true difference. Nevertheless, these analyses replicate the general pattern of head fixation duration on the critical objects, demonstrating that low-experience operators spent less time fixating on the right side and more time on the left side of the simulator on their second drive compared to their first drive.

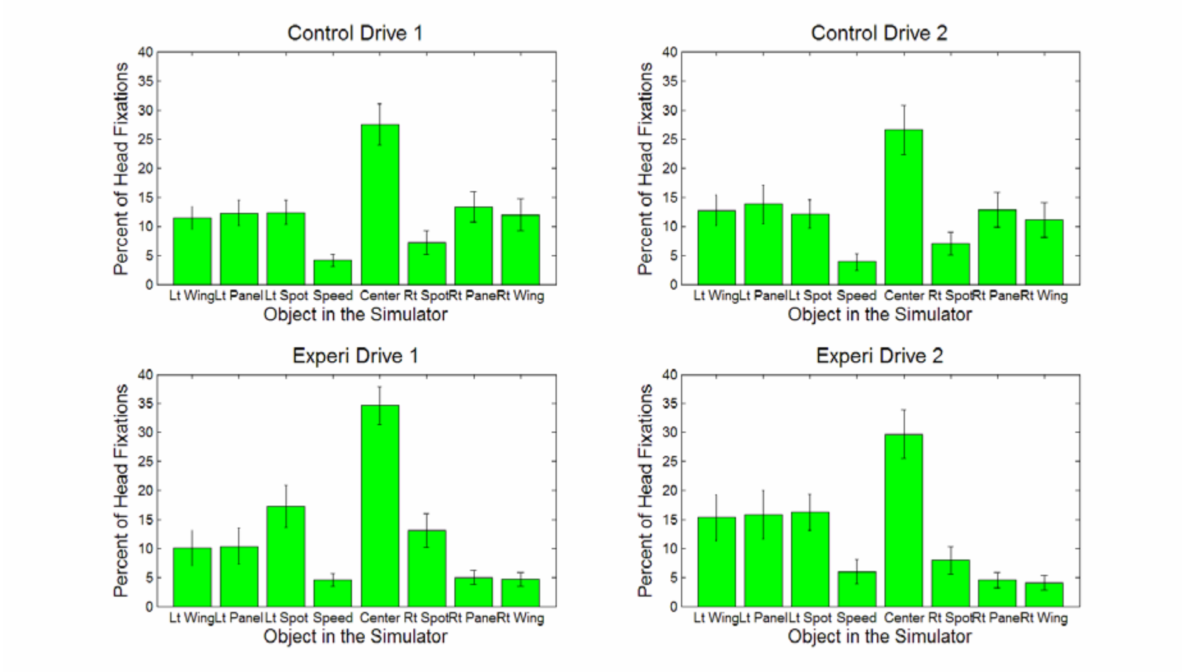


Figure 18. Mean percentage of fixations on each object for low-experience drivers in the control and experimental groups on the first and second drives

To clarify, these results show that low-experience operators in both the control and experimental groups showed a tendency to look less to the right side of the simulator and more to the left side on their second experimental drive compared to their first drive. Since this pattern occurred both for the control and experimental groups (i.e., there were generally no group interactions), these findings can not be attributed to training. Instead, the most likely explanation is that low-experience operators in both groups learned during their first drive that there was an advantage of looking to the left side of the vehicle; thus, they altered their scanning behaviors for their second drive.

Spread of Search

Spread of search is defined as the variation in fixation location across the horizontal and vertical axes. For instance, an operator who spends more time checking his or her mirrors would have a larger spread of horizontal search, as he or she is fixating on objects that are farther away from the center, than an operator who fixates mostly on the center screen, as he or she would focus primarily on a small, restricted area of the screen. Two 3 (experience) X 2 (group) X 2 (time) ANOVAs examining operators' horizontal and vertical spread of search both revealed a main effect of time ($F(1, 113) = 6.34$, MSE < .001, $p < .05$, and $F(1, 113) = 10.91$, MSE = .001, $p = .001$, respectively). Regardless of experience or group, operators showed a reduced spread of search in their second drive compared to their first drive. Overall, however, the size of this difference was relatively small, with at most a .5 of a degree difference between spread of search over the first and second drive (see Figures 19 and 20). Possibly, there may have been a wider spread of horizontal search overall if operators were required to control the plow and wing while operating the simulator.

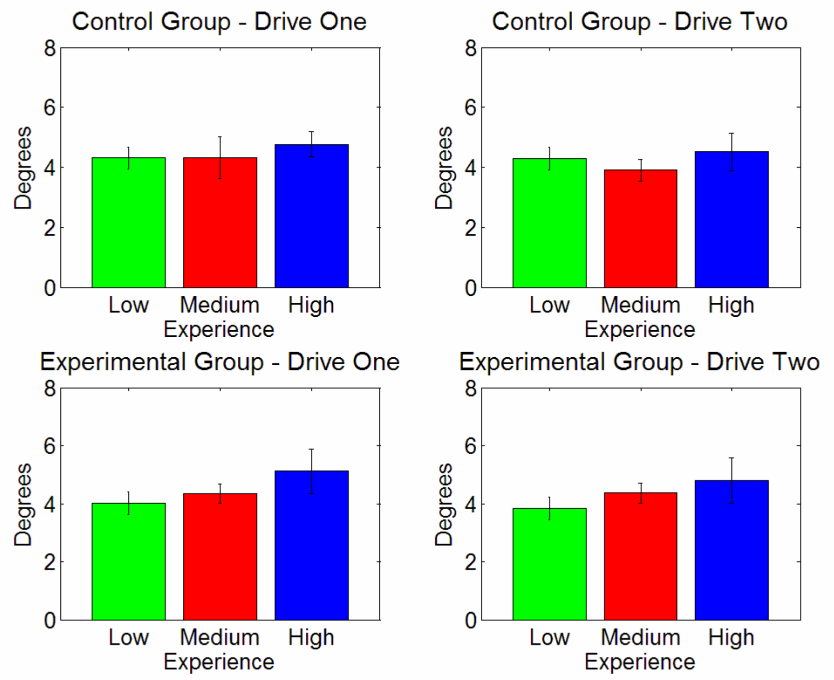


Figure 19. Mean spread of horizontal search for the three experience groups in the control versus experimental conditions during the first and second drives

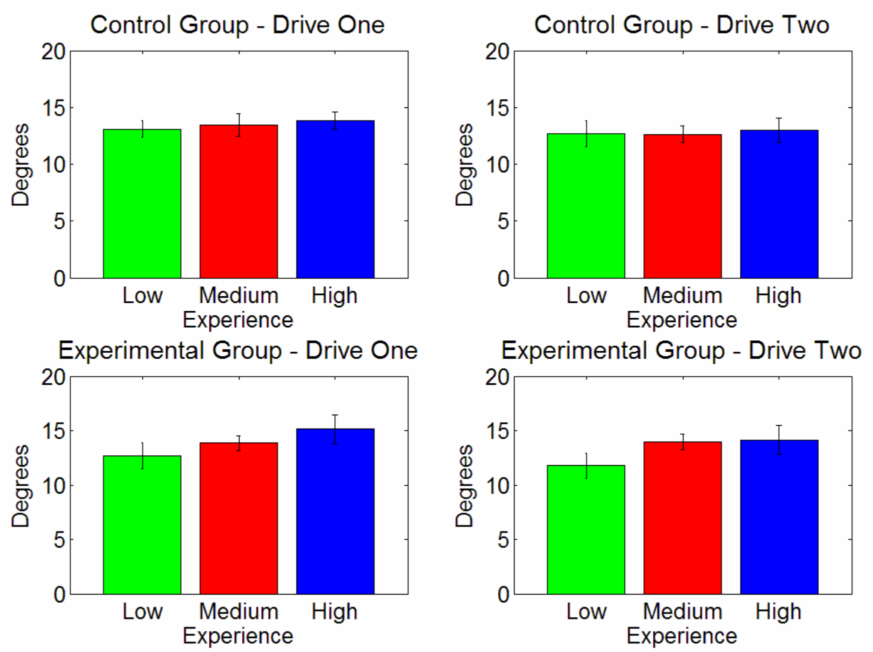


Figure 20. Mean spread of vertical search for the three experience groups in the control versus experimental conditions during the first and second drives

Because this pattern was found in both the control and experimental groups, it is unclear as to what lead to these differences. One possible explanation is that operators adapted to the simulator by their second drive; thus, they may have been behaving similarly to how they would in a real snowplow. Another explanation is that both groups may have learned that there was a benefit for spending more time looking at the center of the road during the first drive in that scenario. Thus, they chose to concentrate more on the center of the roadway during their second drive.

Eye Closure

Research on eye movements suggests that more blinking or eye closure is typically associated with boredom or less effort. Specifically, more blinking or larger eye closures is taken as evidence that participants are comfortable or exerting less effort on the task. A pair of 3 (experience) X 2 (group) X 2 (time) ANOVAs were conducted on participants' blink frequency, as well as on their overall percentage of eye closure. The analysis of blink frequency revealed a main effect of time ($F(1, 63) = 4.97$, MSE =.005, $p < .05$), suggesting that drivers tended to blink more during their second drive. Additionally, while the main effect of experience did not reach statistical significance (p=.12), an LSD comparison showed that high-experience operators tended to blink more often than low-experience operators (see Figure 21).

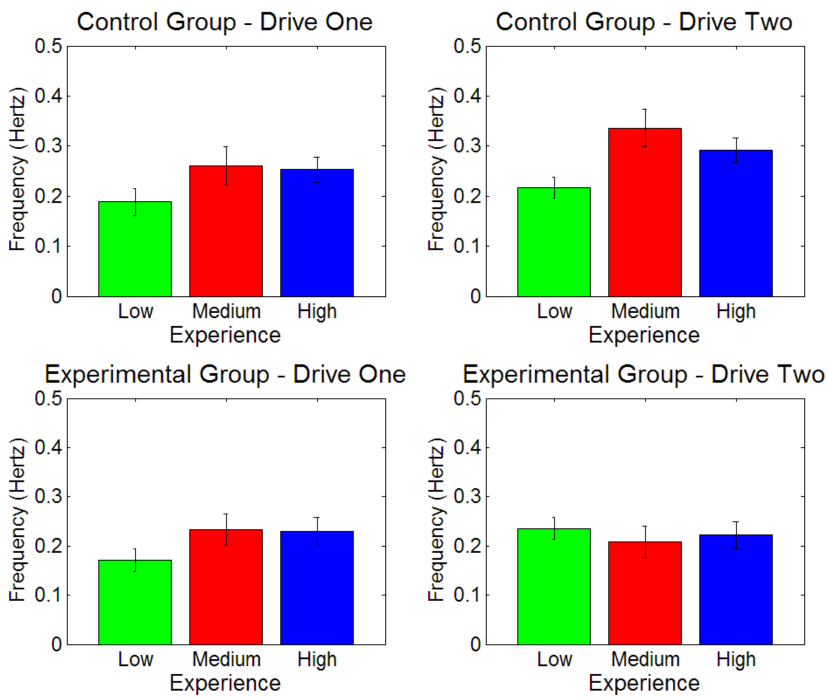


Figure 21. Mean blink frequency for the three experience groups in the control versus experimental conditions during the first and second drives

The results from the analysis of eye closure converged with these findings. The ANOVA revealed a trend for a main effect of time ($F(1, 63) = 3.56$, MSE<.001, $p = .064$), with drivers tending to have their eyes closed more during their second drive than during their first drive. Moreover, although the main effect of experience once again failed to reach statistical significance (p = .135), an LSD comparison showed that high-experience participants had their eyes closed more than medium-experience operators (see Figure 22). Taken together, these results suggest that operators may have found their second drive less difficult than their first, and that the simulator driving may have been less demanding for highly experienced operators.

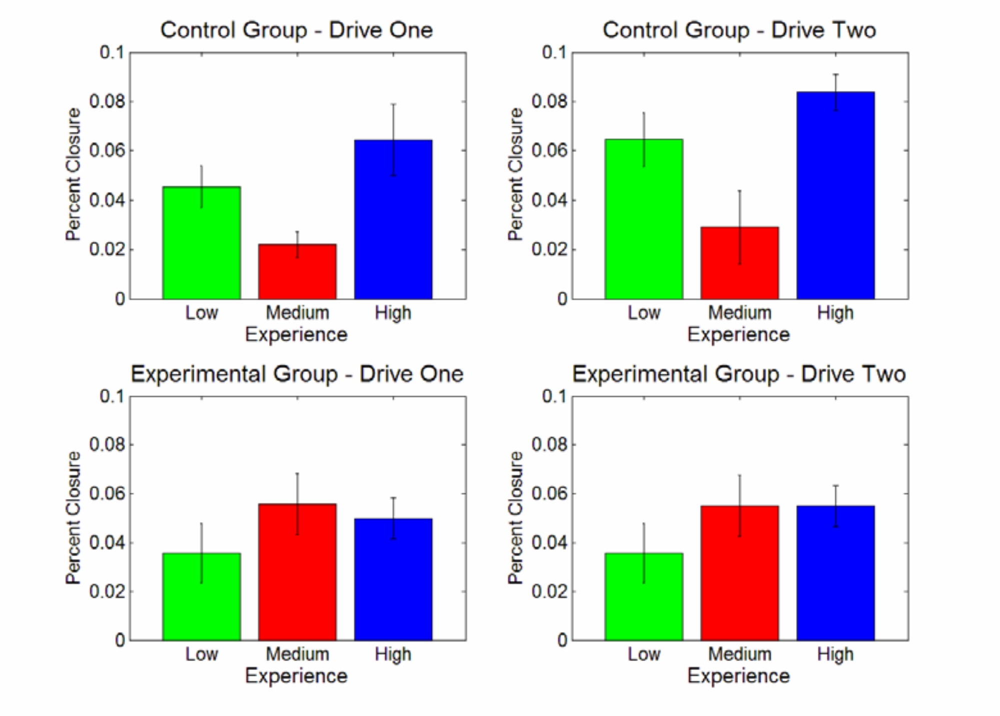


Figure 22. Mean percentage of eye closure for the three experience groups in the control versus experimental conditions during the first and second drives

Summary of Tracking Results

The analyses of object fixation time and fixation changes showed that operators in the low-experience group demonstrated a pattern of looking more towards the left side of the simulator on their second drive and a reduction in the amount of time spent looking towards the right side. Since this trend was found for operators in the experimental and control groups, the most likely explanation is that these drivers learned that there was a benefit of looking towards the left side of the simulator. The spread of search analyses showed that operators tended to have wider horizontal and vertical scan paths during their first drive compared to their second drive. However, this difference was quite small, and may have been due to operators adopting a more narrow searching strategy due to already driving the scenario and thus knowing where many of the critical events were. Finally, high-experience operators showed a tendency to have their eyes closed more than other drivers, implying that the task of driving in the simulator was not as strenuous for them.

Taken as a whole, the tracking data suggest the following:

- Low-experience drivers spent more time looking at the left side of the simulator and less time looking at the right side of the simulator during their second experimental drive compared to their first.
- Although differences in the width of horizontal and vertical scan paths were found for operators' first and second drives, they were quite small and may have occurred by chance.
- The pattern of eye closures is consistent with the interpretation that experienced operators found driving the simulator less demanding than low-experience operators.

Summary of Results

Overall, we found few personality or dispositional differences between less experienced and more experienced snowplow operators. The differences we found typically disappeared once age was taken into account by treating it as a covariant. Thus, personality variables, measured by the NEO-FFI, Zuckerman's (1979) sensation seeking questionnaire, or an immersion questionnaire (Whitmer & Singer 1994), did not appear to differ as a function of experience. There did appear to be some differences between snowplow operators' responses to the NEO-FFI questionnaire and the sensation seeking questionnaire compared to studies that report scores for a normative American male population. However, it is premature to make any decisions based on these data without first comparing the responses of snowplow operators to a more appropriate sample, such as non–snowplow operating Iowa DOT employees.

Additionally, overall levels of reported simulator sickness were modest. Based on the scale used in the questionnaire, the mean responses for the categories measured were below the level of response of slight discomfort for all symptoms. Realism measurements were overall high, as operators reported that seven out of the eight features investigated by the questionnaire were above a realism rating of neutral. Finally, operators' reactions to training were overall quite positive. Many operators reported in their qualitative comments that the driving simulator training was the most useful aspect of training, and the lecture and computer-based training were rated highly also.

Measures of operators' performance in the simulator during the two experimental drives indicated that less experienced drivers tended to be involved in more collisions than the other two groups of more experienced drivers. The most likely explanation for this finding is that experienced drivers managed to apply the skills they learned in the field to operating the simulator. Moreover, low-experience drivers may have shown more improvement as a result of training, but difficulties interpreting the results of the performance output files make drawing any conclusions difficult. Finally, drivers from all experience levels in the experimental group drove faster during their second drive while consuming a comparable amount of fuel than the control group operators. Thus, the driving simulator may be a practical tool for teaching drivers better fuel management techniques.

Finally, there were some fixation differences for operators in the low-experience group between their first and second drives in the simulator. Specifically, they reduced the amount of time looking towards the right side, and increased the amount of time they spent looking towards the left side of the simulator. Also, there did appear to be some evidence that drivers adopted a more narrow spread of search for their second drive compared to their first drive regardless of whether they received training or not. One interpretation of this finding is that by the second drive the majority of participants had adapted to the simulator and thus their fixation behaviors may have been mimicking how they would scan while driving a real snowplow; although, since this difference was so small, it may have simply occurred by chance. Additionally, more experienced drivers tended to blink more often than less experienced drivers, which in the driving research literature is taken as an indication that they were not concentrating as hard, or the task was not as strenuous for them. Thus, converging with the evidence on collision frequency, these results suggest that more experienced operators had an easier time operating the simulator than less experienced operators.

CONCLUSIONS AND INFERENCES

The results of this study suggest that Iowa DOT operators enjoyed and seemed to benefit from virtual reality snowplow simulator training. Operators from all age groups and levels of experience reported having similar immersive tendencies in their everyday life and experienced a similar amount of presence within the simulator. We interpret this to mean that operators of all ages and levels of experience should have the potential to benefit from training equally well. The responses to the training questionnaire tend to support the following explanation: operators from all three levels of experience rated all aspects of training similarly. Additionally, although there was a lot of variation in reported simulator sickness scores, mean ratings were relatively low. There was some evidence that amounts of simulator sickness increased with age, although the mean ratings for all experience groups were less than what would constitute a response of slight for the listed symptoms. These ratings will likely be even lower for subsequent training regimens: more time can be allotted for drivers to become acclimated with the simulator, as this schedule required additional time for the experimental protocol.

The most positively rated aspect of training was the lecture portion. Operators also rated the quality of the trainers very highly, suggesting that they appreciated and learned well from peer instruction. The other two components of training, computer training and simulator training, were rated almost equally and were generally well received. Approximately 80% of participants were not opposed to simulator or computer training, and many reported on the free response section that the simulator portion was the most useful aspect of training. Moreover, many of the responses to the free response questions elucidated what operators felt was lacking from simulator training. These comments will be discussed shortly.

The performance data suggest that there were some noticeable benefits of training. Specifically, low-experience operators who received training prior to their second experimental drive may have shown an improvement in the number of collisions they accrued from their first drive than low-experience operators who did not receive training before their second experimental drive. Difficulties in interpreting the performance files from the simulator software, however, make analyzing these results difficult. This finding was actually somewhat unexpected, considering that training only lasted for around 1.5 to 2 hours. Training programs for less experienced drivers that are able to devote more time to training, particularly within the driving simulator, would have a better chance of demonstrating performance improvements.

Operators in the experimental group also showed a significant increase in their average speed in their second drive compared to their first drive, as well as compared to the second drive for operators in the control group. Importantly, this increase in speed came with a negligible increase in fuel consumption. Thus, this finding also shows that drivers who received training tended to perform better in the simulator than drivers who did not receive training before their second experimental drive.

Low-experience operators seemed to intuitively adapt to driving the simulator. Specifically, operators in the control and experimental groups spent more time looking at the left panel and left wing mirror and less time looking at the right panel and right spot mirror on their second drive compared to their first. Possibly, they may have concluded during the first drive that there was a benefit of looking at the left side of the simulator as opposed to the right, or they may have become more comfortable with the virtual reality environment and started behaving as they normally do when driving a snowplow.

Finally, there were some eye closure differences between operators as a function of experience. More experienced operators demonstrated tendencies associated with being more comfortable or not needing to exert as much effort on driving. Specifically, they tended to show a higher blink frequency and more eye closure than low-experience operators. Both of these behaviors have been interpreted to mean that the driver is not concentrating as hard on the task.

Thus, there are several indications that drivers are behaving in the simulator as one would expect based on their amount of snowplowing experience. Less experienced operators showed a tendency to be involved in more accidents than more experienced operators, while experienced operators seemed not to concentrate as hard on driving in the simulator. Training also seemed to benefit novice operators in particular, and there is some evidence that all drivers who received training used better driving habits on their second experimental drive.

The major caveat of this study, though, is that these results only directly apply to behaviors within the simulator. Although drivers who received training appeared to perform better on their second drive than those who had not had training at that point, it is unclear how these benefits would transfer to real-world snowplowing behaviors. Previous research has shown that individuals tend to behave similarly in driving simulators as they do in real-world settings (Godley, Triggs, and Fildes 2002; Panerai et al. 2001; Tornros 1998), although this has never directly been tested for snowplowing. Another potential concern is that the benefits for the experimental drivers may simply reflect the fact that these operators had additional time to drive the simulator during training. This explanation is actually unlikely as drivers only drove two 5- to 10-minute scenarios during training, which involved truck driving rather than snowplow operating. Thus, the benefits seen as a result of training were not likely due solely to the fact that participants in the experimental group spent more time in the simulator. Information learned during the lecture and computer-based portions of training were undoubtedly influential on improving operators' performance as well.

There are also some other minor concerns with the experimental paradigm. First, several different trainers were used throughout this study. Although all trainers used the same training scenarios and followed a predetermined script for the lecture, operators may have responded to some trainers more positively than to others. Potentially, this may have increased the variance in participants' data, thus reducing the likelihood of finding statistically significant results. Additionally, the simulator upgrade at the end of the third week of training excluded approximately 35 participants from being included in many of the performance analyses. Several of the thresholds for causing certain types of failure were altered during the upgrade; thus, it would have been misleading to include these operators in the statistical analysis for these performance measures. These participants were, however, included in the questionnaire, fixation, and collision measures, as the minor improvements made during the upgrade were not likely to affect any of these measures.

It is important to keep in mind, also, that each session lasted only four hours, and approximately half of this time was allocated to tasks involving this experiment. Training itself only lasted on average from 1.5 to 2 hours, and as mentioned above, operators were only scheduled for two trained drives in the simulator, which even involved non-snowplow scenarios. Although the purpose of training was to disseminate better driving practices rather than to teach operators how to plow snow per se, allocating more time for training would increase the chances that skills learned during training would generalize to real-world performance.

There are certainly several methods to judge the effectiveness or generalizability of simulator training in real-world situations. One possibility would be to compare the performance, during one or more winter seasons, of a group of operators who received training versus a matched control group who did not receive training. However, this would require a large portion of drivers to forgo what may be highly valuable training. Thus, a possible alternative would be to train as many drivers as possible, even bringing operators back for multiple sessions if possible, and compare overall collision or fuel-consumption rates for the next few years to an equal number of previous years. Although this method lacks complete experimental control, it allows for a greater number of operators to receive training and should provide a critical examination of the effectiveness of this particular training regimen in real-world situations. Conversely, short-term tests are also an option. For instance, a group of trained operators and a matched control group can be tested on a series of courses designed to approximate the demands of operating a snowplow. Differences between the groups' performance could then be attributed to whether or not they received training. This option may not be ideal, though, as there would still be concern over the generalizability of those results to actual snowplow operation in the field.

Another area of interest may be to improve the simulator itself based on participants' recommendations. Generally, the simulator reality ratings were consistent with participants' free response comments. For instance, one of the lower realism ratings was for the responsiveness of the controls. Accordingly, many operators noted that (a) the relationship between engine sound and speed was incongruous, (b) steering was too sensitive, and (c) the vehicle tended to accelerate and slow down unrealistically fast. Several operators also commented that driving conditions during storms were not realistically captured by the simulator. For instance, they noted that visibility is typically poorer, and there is usually more snow and slush on the windshield. All of these suggestions, so far, are easily addressable. Many operators also suggested that the simulator would be a much more effective training tool if the driver had control over the wing and plow, and if features such as a salt spreader were included. Many of these potential upgrades reported here are currently being added or have already been implemented in the snowplow simulator.

Another interesting finding, which may warrant some additional research, was that participants reported some differences across the five NEO-FFI personality subscales compared to a population of normative American males. Additionally, operators also reported less sensation seeking tendencies than the general population, although this may be an artifact of only examining some of the subscales or an overall higher mean age in this sample compared to the normative group. To make any strong inferences about personality differences between Iowa DOT operators that may be characteristic of becoming a successful snowplow operator, it would be necessary to identify a more suitable control group—for instance, Iowa DOT employees who do not operate snowplows. This would control for important factors such as age, race, and socioeconomic class that may have been responsible for the differences that we found.

Overall, this study found the following:

- On the basis of respondents immersion and presence scores, operators of all levels of experience should be able to benefit from virtual reality snowplow simulator training
- Simulator sickness ratings were relatively low, implying that the simulator is appropriate for training a wide range of Iowa DOT operators
- Iowa DOT operators tended to enjoy training, and many reported that simulator training was the most useful aspect of training for them
- There is some evidence that operators who received training prior to completing their second drive in the simulator performed better than operators in the control group
- Snowplow experience was associated with performance, as more experienced operators tended to perform better in the simulator
- Low-experience operators showed a tendency to alter their fixation behaviors from the first to the second drive.

REFERENCES

Barrick, M. R., and M. K. Mount. 1991. The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology* 44: 1-27.

Campagne, A., T. Pebayle, and A. Muzet. 2005. Oculomotor changes due to road events during prolonged monotonous simulated driving. *Biological Psychology* 68:353–368.

Costa, P. T., and R. R. McCrae. 1985. *The NEO Personality Inventory Manual*. Odessa, FL: Psychological Assessment Resources.

Duh, H. B. L., D. E. Parker, and T. A. Furness. 2004. An independent visual background reduced simulator sickness in a driving simulator. *Presence* 13:578–588.

Godley, S. T., T. J. Triggs, and B. N. Fildes. 2002. Driving simulator validation for speed research. *Accident Analysis and Prevention* 24:589–600.

Horberry, T., J. Anderson, and M. A. Regan. 2005. The possible safety benefits of enhanced road markings: A driving simulator evaluation. *Transportation Research Part F: Traffic Psychology and Behaviour* 9:77–87.

Kemeny, A., and F. Panerai. 2003. Evaluating perception in driving simulation experiments. *Trends in Cognitive Sciences* 7:31–37.

Kennedy, R. S., N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. 1993. A simulator sickness questionnaire (SSQ): A new method for quantifying simulator sickness. *International Journal of Aviation Psychology* 3:203–220.

LaViola, J. J. 2000. A discussion of cybersickness in virtual environments. *SIGCHI Bulletin* 32:47–56.

Masciocchi, C. M., V. J. Dark, and D. Parkhurst. 2006. Evaluation of Virtual Reality Snowplow Simulator Training: A Literature Review. CTRE Project 06-245.

Miller, T. R. 1992. Benefit-cost analysis of lane marking. *Transportation Research Record* 1334:38–45.

Panerai, F., J. Droulez, J. M. Kelada, A. Kemeny, E. Balligand, and B. Favre. 2001. Speed and safety distance control in truck driving: Comparison of simulation and real-world environment. *DSC2001–Sophia Antipolis.*

Rolland, J. P., W. D. Parker, and H. Stumpf. 1998. A psychometric examination of the French translations of the NEO-PI-R and NEO-FFI. *Journal of Personality Assessment* 71:269–291.

Tett, R. P., D. N. Jackson, and M. Rothstein. 1991. Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology* 44: 703-742.

Tornros, J. 1998. Driving behavior in a real and a simulated road tunnel: A validation study. *Accident Anaysis and Prevention* 30:497–503.

Santos, J., N. Merat, S. Mouta, K. Brookhuis, and D. Waard. 2005. The interaction between driving and in-vehicle information systems: Comparison of results from laboratory, simulator and real-world studies. *Transportation Research Part F: Traffic Psychology and Behaviour* 8:135–146.

Sodhi, M., B. Reimer, J. L. Cohen, E. Vastenburg, R. Kaars, and S. Kirschenbaum. 2002. On-road driver eye movement tracking using head-mounted devices. Paper presented at the Eye Tracking Research and Applications Symposium, New Orleans, LA.

Stanney, K. M., and G. Salvendy. 1998. Aftereffects and sense of presence in virtual environments: Formulation of a research and development agenda. *International Journal of Human-Computer Interaction* 10:135–187.

Strayer, D. L., F. A. Drews, and S. Burns. 2004. The development and evaluation of a high-fidelity simulator training program for snowplow operators. Utah Department of Transportation Report UT-04.17.

Underwood, G, D. Crundall, and P. Chapman. 2002. Selective searching while driving: The role of experience in hazard detection and general surveillance. *Ergonomics* 45:1–12.

Witmer, B. G., and M. J. Singer. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence* 7:225–240.

Zuckerman, M. 1979. *Sensation Seeking: Beyond the Optimal Level of Arousal.* Hillsdale, NJ: Erlbaum.

APPENDIX A: COPY OF LECTURE POWER POINT

L3 MPRI

101

Circles of Influence

L3 MPRI

Learning Objectives

- List the Circles of Influence that affect driving performance and decision-making.
- Define the steps in the SIPDE Method to hazard perception and apply each to a driving situation.
- Demonstrate the ability to make successful decisions in simulation using applicable aspects of the Circles of Influence methodology.

L3 MPRI

What Is the Leading Cause of Collisions?

- Poor vehicle-handling skills
- Poor decision-making skills

80%

L3 MPRI

The Decision-Making Process

- Preparation
- Practice

L3 MPRI

Driver Safety

- Most driver-safety programs target the 15% of drivers who get in collisions.
- Management strategies should be designed to help the 85% of drivers who drive safely.

L3 MPRI

The Circles of Influence

- You
- Health and family
- Job and equipment knowledge
- Environment
- Hazard perception
- Successful decision-making

You

- Age, gender, personality, and region you're from
- Driving experience and familiarity with equipment
- Traffic violations and collisions

Health and Family

- Breaks, rest, relaxation, and exercise
- Family crises and unresolved issues
- Fatigue
- Medications

Job and Equipment Knowledge

- DOT policies
- Pre-trip inspections
- Vehicle and equipment
- Special gear or paperwork

Environmental Conditions

- Weather
- Traffic
- Road conditions and structures

Hazard Perception—The SIPDE Method

- **S**can and **S**earch
- **I**dentify
- **P**redict
- **D**ecide
- **E**xecute

The SIPDE Method

***S**can and **S**earch:*

- Search for hazards
- Avoid tunnel vision
- Be alert to sounds, smells, and vibrations

MPRI

The SIPDE Method

Identify:

- Interpret sensory input
- Sort information
- Identify hazards
- Weigh dangers

MPRI

The SIPDE Method

Predict:

- Imagine result of potential threat
- Decide if you can change the result
- Anticipate what other drivers might do

MPRI

The SIPDE Method

Decide:

- Critical moments often require fast decisions
- Practice decision-making by imagining conflicts and outcomes

MPRI

The SIPDE Method

Execute:

- Quickly follow decisions with action
- Maintain strong vehicle-handling skills

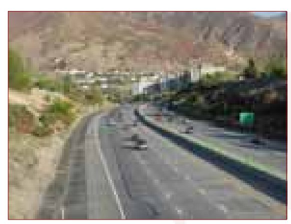
MPRI

Benefits of Successful Decision-Making

- Job satisfaction and self-esteem
- Confidence and trust
- More good decisions
- Wise driving habits

MPRI

Things to think about !



MPRI

Maintaining Focus

- Talk out loud.
- Analyze what other drivers are doing.

MPRI

Driving Consistently

- Hold your lane, speed, and position.
- Your consistency helps others make safe decisions.

MPRI

On to the Simulator!

Advanced Simulation Exercises
TTSCN_04

MPRI

402
Space Management

MPRI

Learning Objectives

- Calculate a safe following distance under a variety of circumstances by applying the following-distance formula.
- List appropriate space-management techniques used to manage the space around a vehicle when encountering a variety of hazards.
- Demonstrate ability to manage space cushion around a vehicle when encountering typical driving hazards in simulation.

MPRI

CBT Lab
Space Management Part 1

5 Minute Break

On to the CBT Lab!
Space Management Part 2

SIMULATOR
First Driver: TTR 402C
Second Driver: TTSCN_8

SUMMARY

Key Points: Circles of Influence

- You
- Health and family
- Knowledge
- Environment
- Hazard perception
- Successful decision-making

Key Points: SIPDE Method ?

- **S**can and **S**earch
- **I**dentify
- **P**redict
- **D**ecide
- **E**xecute

Key Points: Space, Stopping Time

- The space cushion: space ahead, behind, above, below, and to the sides.
- The timed-interval method indicates the following distance in seconds.
- Stopping time includes time to see, think, react, and brake.
- Calculate stopping distance according to vehicle length. Adjust for adverse conditions.

Key Points: Best Practices

- Drivers develop their own best practices through experience.
- Drivers benefit from one another's experiences when they share best practices.

Key Points: Maneuverability, Right-of-Way

- Maneuverability depends on vehicle size, weight, load, speed, center of gravity, and condition.
- Managing space during turns, curves, passes, merges, and lane changes is critical to safe driving.
- Respect laws for yielding right-of-way.

APPENDIX B: NEO-FFI QUESTIONNAIRE

Please rate the extent to which you agree with the following statements.
Use the scale below for your responses.

Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Agree
1	2	3	4	5

______ 1. I am not a worrier.
______ 2. I like to have a lot of people around me.
______ 3. I don't like to waste my time daydreaming.
______ 4. I try to be courteous to everyone I meet.
______ 5. I keep my belongings neat and clean.
______ 6. I often feel inferior to others.
______ 7. I laugh easily.
______ 8. Once I find the right way to do something, I stick to it.
______ 9. I often get into arguments with my family and co-workers.
______ 10. I'm pretty good about pacing myself so as to get things done on time.
______ 11. When I'm under a great deal of stress, sometimes I feel like I'm going to pieces.
______ 12. I don't consider myself especially "light-hearted."
______ 13. I am intrigued by the patterns I find in art and nature.
______ 14. Some people think I'm selfish and egotistical.
______ 15. I am not a very methodical person.
______ 16. In dealing with other people, I always dread making a social blunder.
______ 17. I really enjoy talking to people.
______ 18. I believe letting students hear controversial speakers can only confuse and mislead them.
______ 19. I would rather cooperate with others than compete with them.
______ 20. I try to perform all the tasks assigned to me conscientiously.
______ 21. I often feel tense and jittery.
______ 22. I like to be where the action is.
______ 23. Poetry has little or no effect on me.
______ 24. I tend to be cynical and skeptical of others' intentions.
______ 25. I have a clear set of goals and work toward them in an orderly fashion.
______ 26. Sometimes I feel completely worthless.
______ 27. I usually prefer to do things alone.
______ 28. I often try new and foreign foods.
______ 29. I believe that most people will take advantage of you if you let them.
______ 30. I waste a lot of time before settling down to work.
______ 31. I rarely feel fearful or anxious.
______ 32. I often feel as if I'm bursting with energy.
______ 33. I seldom notice the moods or feelings that different environments

produce.

______ 34. Most people I know like me.
______ 35. I work hard to accomplish my goals.
______ 36. I often get angry at the way people treat me.
______ 37. I am a cheerful, high-spirited person.
______ 38. I believe we should look to our religious authorities for decisions on moral issues.
______ 39. Some people think of me as cold and calculating.
______ 40. When I make a commitment, I can always be counted on to follow through.
______ 41. Too often, when things go wrong, I get discouraged and feel like giving up.
______ 42. I am not a cheerful optimist.
______ 43. Sometimes when I am reading poetry or looking at a work of art, I feel a chill or wave of excitement.
______ 44. I'm hardheaded and tough-minded in my attitudes.
______ 45. Sometimes I'm not as dependable or reliable as I should be.
______ 46. I am seldom sad or depressed.
______ 47. My life is fast-paced.
______ 48. I have little interest in speculating on the nature of the universe or the human condition.
______ 49. I generally try to be thoughtful and considerate.
______ 50. I am a productive person who always gets the job done.
______ 51. I often feel helpless and want someone else to solve my problems.
______ 52. I am a very active person.
______ 53. I have a lot of intellectual curiosity.
______ 54. If I don't like people, I let them know it.
______ 55. I never seem to be able to get organized.
______ 56. At times I have been so ashamed I just wanted to hide.
______ 57. I would rather go my own way than be a leader of others.
______ 58. I often enjoy playing with theories or abstract ideas.
______ 59. If necessary, I am willing to manipulate people to get what I want.
______ 60. I strive for excellence in everything I do.

APPENDIX C: SENSATION SEEKING QUESTIONNAIRE

Interest and Preference Test

Each of the items below contains two choices, A and B. Please indicate which of the choices most describes your likes or the way you feel. In some cases you may find items in which both choices describe your likes or feelings. Please choose the one which better describes your likes or feelings. In some cases you may find items in which you do not like either choice. In these cases mark the choice you dislike least. Do not leave any items blank. It is important you respond to all items with only one choice, A or B. We are interested only in your likes or feelings, not in how others feel about these things or how one is supposed to feel. There are no right or wrong answers as in other kinds of tests. Be frank and give your honest appraisal of yourself.

1. A. I like uninhibited parties.
   B. I prefer quiet parties with good conversation.

2. A. There are some movies I enjoy seeing a second or even a third time.
   B. I can’t stand watching a movie that I’ve seen before.

3. A. I often wish I could be a mountain climber.
   B. I can’t understand people who risk their necks climbing mountains.

4. A. I dislike all body odors.
   B. I like some of the earthly body smells.

5. A. I get bored seeing the same old faces.
   B. I like the comfortable familiarity of everyday friends.

6. A. I like to explore a strange city or section of town by myself, even if it means getting lost.
   B. I prefer a guide when I am in a place I don’t know well.

7. A. I dislike people who do or say things just to shock or upset people.
   B. When you can predict almost everything a person will do and say he or she must be a bore.

8. A. I usually don’t enjoy a movie or play where I can predict what will happen in advance.
   B. I don’t mind watching a movie or play where I can predict what will happen in advance.

9 A. A sensible person avoids activities that are dangerous.
   B. I sometimes like to do things that are a little frightening.

10. A. I like to try new foods that I have never tasted before.
    B. I order the dishes with which I am familiar so as to avoid disappointment and unpleasantness.

11. A. I enjoy looking at home movies, videos, or travel slides.
    B. Looking at someone’s home movies, videos or travel slides bores me tremendously.

12. A. I would like to take up the sport of water skiing.
    B. I would not like to take up water skiing.

13. A. I would like to try surfing.
B. I would not like to try surfing.

14. A. I would like to take off on a trip with no preplanned or definite routes, or timetable.
B. When I go on a trip I like to plan my route and timetable fairly carefully.

15. A. I prefer the "down to earth" kinds of people as friends.
B. I would like to make unusual friends.

16. A. I would not like to learn to fly an airplane.
B. I would like to learn to fly an airplane.

17. A. I prefer the surface of the water to the depths.
B. I would like to go scuba diving.

18. A. I would like to try parachute jumping.
B. I would never want to try jumping out of a plane, with or without a parachute.

19. A. I prefer friends who are excitingly unpredictable.
B. I prefer friends who are reliable and predictable.

20. A. I am not interested in experience for its own sake.
B. I like to have new and exciting experiences and sensations even if they are a little frightening or unconventional.

21. A. The essence of good art is in its clarity, symmetry of form, and harmony of colors.
B. I often find beauty in the "clashing" colors and irregular forms of modern painting.

22. A. I enjoy spending time in the familiar surroundings of home.
B. I get very restless if I have to stay around home for any length of time.

23. A. I like to dive off the high board.
B. I don't like the feeling I get standing on the high board (or I don't go near it at all).

24. A. I like to date persons who are physically exciting.
B. I like to date persons who share my values.

25. A. The worst social sin is to be rude.
B. The worst social sin is to be a bore.

26. A. I like people who are sharp and witty even if they do sometimes insult each other.
B. I dislike people who have their fun at the expense of hurting the feelings of others.

27. A. People should dress according to some standard of taste, neatness and style.
B. People should dress in individual ways even if the effects are sometimes strange.

28. A. Sailing long distances in small sailing crafts is foolhardy.
B. I would like to sail a long distance in a small but seaworthy sailing craft.

29. A. I have no patience with dull or boring people.
B. I find something interesting in almost every person I talk to.

30. A. Skiing down a high mountain slope is a good way to end up on crutches.
B. I think I would enjoy the sensations of skiing very fast down a high mountain slope.

APPENDIX D: IMMERSION QUESTIONNAIRE

Pre experiment questionnaire
Please read the questions carefully and circle the appropriate number from 1 to 7 .

1. Do you ever get extremely involved in projects that are assigned to you by your boss or your instructor, to the exclusion of other tasks?

1	2	3	4	5	6	7
NOT OFTEN			MODERATELY OFTEN			VERY OFTEN

2. How easily can you switch your attention from the task in which you are currently involved to a new task?

1	2	3	4	5	6	7
NOT EASILY			MODERATELY EASILY			VERY EASILY

3. How frequently do you get emotionally involved (angry, sad, or happy) in the news stories that you read or hear?

1	2	3	4	5	6	7
NOT FREQUENTLY			MODERATELY FREQUENTLY			VERY FREQUENTLY

4. How well do you feel today?

1	2	3	4	5	6	7
NOT WELL			MODERATELY WELL			VERY WELL

5. Do you easily become deeply involved in movies or TV dramas?

1	2	3	4	5	6	7
NOT EASILY			MODERATELY EASILY			VERY EASILY

6. Do you ever become so involved in a television program or book that people have problems getting your attention?

1	2	3	4	5	6	7
NOT OFTEN			MODERATELY OFTEN			VERY OFTEN

7. How mentally alert do you feel at the present time?

1	2	3	4	5	6	7
NOT ALERT			MODERATELY ALERT			VERY ALERT

8. Do you ever become so involved in a movie that you are not aware of things happening around you?

1	2	3	4	5	6	7
NOT OFTEN			MODERATELY OFTEN			VERY OFTEN

9. How frequently do you find yourself closely identifying with the characters in a story line?

1	2	3	4	5	6	7
NOT OFTEN			MODERATELY OFTEN			VERY OFTEN

10. Do you ever become so involved in a video game that it is as if you are inside the game rather than moving a joystick and watching the screen?

1	2	3	4	5	6	7
NOT OFTEN			MODERATELY OFTEN			VERY OFTEN

11. On average, how many books do you read for enjoyment in a month?

12. What kind of books do you read most frequently?
(CIRCLE ONE ITEM ONLY!)

Spy novels Fantasies Science fiction Adventure

Romance novels Historical novels
Westerns Mysteries Other fiction Biographies

Autobiographies Other non-fiction

13. How physically fit do you feel today?

1	2	3	4	5	6	7
NOT FIT			MODERATELY FIT			VERY FIT

14. How good are you at blocking out external distractions when you are involved in something?

1	2	3	4	5	6	7
NOT GOOD			MODERATELY			VERY GOOD

15. When watching sports, do you ever become so involved in the game that you react as if you were one of the players?

1	2	3	4	5	6	7
NOT OFTEN			MODERATELY OFTEN			VERY OFTEN

16. Do you ever become so involved in a daydream that you are not aware of things happening around you?

1	2	3	4	5	6	7
NOT OFTEN			MODERATELY OFTEN			VERY OFTEN

17. Do you ever have dreams that are so real that you feel disoriented when you awake?

1	2	3	4	5	6	7
NOT OFTEN			MODERATELY OFTEN			VERY OFTEN

18. When playing sports, do you become so involved in the game that you lose track of time?

1	2	3	4	5	6	7
NOT OFTEN			MODERATELY OFTEN			VERY OFTEN

19. Are you easily disturbed when working on a task?

1	2	3	4	5	6	7
NOT EASILY			MODERATELY EASILY			VERY EASILY

20. How well do you concentrate on enjoyable activities?

1	2	3	4	5	6	7
NOT WELL			MODERATELY WELL			VERY WELL

21. How often do you play arcade or video games? (OFTEN should be taken to mean every day or every two days, on average.)

1	2	3	4	5	6	7
NOT OFTEN			MODERATELY OFTEN			VERY OFTEN

22. How well do you concentrate on disagreeable tasks?

1	2	3	4	5	6	7
NOT WELL			MODERATELY WELL			VERY WELL

23. Have you ever gotten excited during a chase or fight scene on TV or in the movies?

1	2	3	4	5	6	7
NOT OFTEN			MODERATELY OFTEN			VERY OFTEN

24. To what extent have you dwelled on personal problems in the last 48 hours?

1	2	3	4	5	6	7
NOT MUCH			MODERATELY			VERY MUCH

25. Have you ever gotten scared by something happening on a TV show or in a movie?

1	2	3	4	5	6	7
NOT OFTEN			MODERATELY OFTEN			VERY OFTEN

26. Have you ever remained apprehensive or fearful long after watching a scary movie?

1	2	3	4	5	6	7
NOT OFTEN			MODERATELY OFTEN			VERY OFTEN

27. Do you ever avoid carnival or fairground rides because they are too scary?

1	2	3	4	5	6	7
NOT OFTEN			MODERATELY OFTEN			VERY OFTEN

28. How frequently do you watch TV soap operas or docu-dramas?

1	2	3	4	5	6	7
NOT OFTEN			MODERATELY OFTEN			VERY OFTEN

29. Do you ever become so involved in doing something that you lose all track of time?

1	2	3	4	5	6	7
NOT OFTEN			MODERATELY OFTEN			VERY OFTEN

APPENDIX E: PRESENCE QUESTIONNAIRE

Post experiment questionnaire
Please read the questions carefully and circle the appropriate number from 1 to 7 .

1. How much were you able to control events?

1	2	3	4	5	6	7
NOT WELL			MODERATELY WELL			VERY WELL

2. How responsive was the environment to actions that you initiated (or performed)?

1	2	3	4	5	6	7
NOT RESPONSIVE			MODERATELY RESPONSIVE			VERY RESPONSIVE

3. How natural did your interactions with the environment seem?

1	2	3	4	5	6	7
NOT NATURAL			MODERATELY NATURAL			VERY NATURAL

4. How completely were all of your senses engaged?

1	2	3	4	5	6	7
NOT ENGAGED			MODERATELY ENGAGED			VERY ENGAGED

5. How much did the visual aspects of the environment involve you?

1	2	3	4	5	6	7
NOT MUCH			MODERATELY			VERY MUCH

6. How much did the auditory aspects of the environment involve you?

1	2	3	4	5	6	7
NOT MUCH			MODERATELY			VERY MUCH

7. How natural was the mechanism which controlled movement through the environment?

1	2	3	4	5	6	7
NOT NATURAL			MODERATELY NATURAL			VERY NATURAL

8. How aware were you of events occurring in the real world around you?

1	2	3	4	5	6	7
NOT AWARE			MODERATELY AWARE			VERY AWARE

9. How aware were you of your display and control devices?

1	2	3	4	5	6	7
NOT AWARE			MODERATELY AWARE			VERY AWARE

10. How compelling was your sense of objects moving through space?

1	2	3	4	5	6	7
NOT COMPELLING			MODERATELY COMPELLING			VERY COMPELLING

11. How inconsistent or disconnected was the information coming from your various senses?

1	2	3	4	5	6	7
NOT INCONSISTENT			MODERATELY INCONSISTENT			VERY INCONSISTENT

12. How much did your experiences in the virtual environment seem consistent with your real-world experiences?

1	2	3	4	5	6	7
NOT CONSISTENT			MODERATELY CONSISTENT			VERY CONSISTENT

13. Were you able to anticipate what would happen next in response to the actions you performed?

1	2	3	4	5	6	7
NOT WELL			MODERATELY WELL			VERY WELL

14. How completely were you able to actively survey or search the environment using vision?

1	2	3	4	5	6	7
NOT COMPLETELY			MODERATELY			COMPLETELY

15. How well could you identify sounds?

1	2	3	4	5	6	7
NOT WELL			MODERATELY WELL			VERY WELL

16. How well could you localize sounds?

1	2	3	4	5	6	7
NOT WELL			MODERATELY WELL			VERY WELL

17. How well could you actively survey or search the environment using touch?

1	2	3	4	5	6	7
NOT WELL			MODERATELY WELL			VERY WELL

18. How compelling was your sense of moving around inside the virtual environment?

1	2	3	4	5	6	7
NOT COMPELLING			MODERATELY COMPELLING			VERY COMPELLING

19. How closely were you able to examine objects?

1	2	3	4	5	6	7
NOT CLOSELY			MODERATELY CLOSELY			VERY CLOSELY

20. How well could you examine objects from multiple viewpoints?

1	2	3	4	5	6	7
NOT WELL			MODERATELY WELL			VERY WELL

21. How well could you move or manipulate objects in the virtual environment?

1	2	3	4	5	6	7
NOT WELL			MODERATELY WELL			VERY WELL

22. To what degree did you feel confused or disoriented at the beginning of breaks or at the end of the experimental session?

1	2	3	4	5	6	7
NOT DISORIENTED			MODERATELY DISORIENTED			VERY DISORIENTED

23. How involved were you in the virtual environment experience?

1	2	3	4	5	6	7
NOT INVOLVED			MODERATELY INVOLVED			VERY INVOLVED

24. How distracting was the control mechanism?

1	2	3	4	5	6	7
NOT DISTRACTING			MODERATELY DISTRACTING			VERY DISTRACTING

25. How much delay did you experience between your actions and expected outcomes?

1	2	3	4	5	6	7
NOT MUCH			MODERATELY			VERY MUCH

26. How quickly did you adjust to the virtual environment experience?

1	2	3	4	5	6	7
NOT QUICKLY			MODERATELY QUICKLY			VERY QUICKLY

27. How proficient in moving and interacting with the virtual environment did you feel at the end of the experience?

1	2	3	4	5	6	7
NOT PROFICIENT			MODERATELY PROFICIENT			VERY PROFICIENT

28. How much did the the visual display quality interfere or distract you from performing assigned tasks or required activities?

1	2	3	4	5	6	7
NOT MUCH			MODERATELY			VERY MUCH

29. How much did the control devices interfere with performance of assigned tasks or other activities?

1	2	3	4	5	6	7
NOT MUCH			MODERATELY			VERY MUCH

30. How well could you concentrate on the assigned tasks or required activities rather than on the mechanisms used to perform those tasks or activities?

1	2	3	4	5	6	7
NOT WELL			MODERATELY WELL			VERY WELL

31. Did you learn new techniques that enabled you to improve your performance?

1	2	3	4	5	6	7
NOT MUCH			MODERATELY			VERY MUCH

32. Were you involved in the experimental task to the extent that you lost track of time?

1	2	3	4	5	6	7
NOT MUCH			MODERATELY			VERY MUCH

APPENDIX F: SIMULATOR SICKNESS QUESTIONNAIRE

Simulator Sickness Questionnaire

Please report the degree to which you experience each of the below symptoms as one of ``None", ``Slight", ``Moderate" and ``Severe". Using the scale from "0" (none) to "3" (severe).

	None	Slight	Moderate	Severe
General discomfort	0	1	2	3
Fatigue	0	1	2	3
Headache	0	1	2	3
Eyestrain	0	1	2	3
Difficulty focusing	0	1	2	3
Increased salivation	0	1	2	3
Sweating	0	1	2	3
Nausea	0	1	2	3
Difficulty concentrating	0	1	2	3
Fullness of head	0	1	2	3
Blurred vision	0	1	2	3
Dizzy (eyes open)	0	1	2	3
Dizzy (eyes closed)	0	1	2	3
Vertigo	0	1	2	3
Stomach awareness	0	1	2	3
Burping	0	1	2	3

Scoring: Participants report 0, 1, 2 or 3 for each of these above symptoms, corresponding to rating of "none," "slight," "moderate," and "severe." To calculate the total simulator sickness score, multiply each rating by 3.74, and then sum up all 16 values for each subject.

APPENDIX G: TRAINING AND REALISM QUESTIONNAIRE

Please indicate how much you agree or disagree with the following statements.

1) The snowplow training package was very useful

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

2) The classroom/lecture portion of the training was very useful.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

3) The simulator training was very useful.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

4) This training should be part of Iowa DOT training for all snowplow operators.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

5) The training helped prepare me for dealing with non-routine situations.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

6) The training helped prepare me for situations involving passing cars.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

7) The training helped prepare me for situations involving vehicles or pedestrians along the side of the road.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

8) This training explained why speed management is important for safe plowing.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

9) This training explained why space management is important for safe plowing.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

10) This training explained why good communication is important for safe plowing.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

11) I would recommend this training for other snowplow drivers.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

12) The course objectives satisfied my needs.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

13) The driving simulations were realistic for the course objectives.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

14) I practiced skills during the driving simulation part of the course that will be very useful on the road.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

15) The skills I practiced during the SIPDE part of the course will be very useful on the road.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

16) The time spent in the lecture portion of the course was appropriate.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

17) The time spent in the lecture portion of the course was appropriate.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

18) The time spent in the driving simulation portion of the course was appropriate.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

19) The trainer had a good understanding of the course material.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

20) The trainer worked will with the drivers.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

21) The trainer understood your needs and issues.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

22) The trainer gave very useful feedback.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

23) The scenery (trees, sky roads) in the simulator looks realistic.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

24) The vehicles (cars, trucks) in the scenarios look realistic.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

25) The mirrors were where they would be in a real snowplow.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

26) The mirrors worked as well as mirrors in a real snowplow.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

27) The behavior of other vehicles in the scenarios was realistic.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

28) The controls (steering wheel, pedals) in the simulator were where they would be in a real snowplow.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

29) The controls (steering wheel, pedals) in the simulator were as responsive as normal snowplow controls.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree

30) I felt like I was driving a real snowplow while operating the simulator.

1	2	3	4	5	6	7
Strongly Disagree			Neutral			Strongly Agree