# Simulator Sickness Ratings Reduce with Simulator Motion when Driven Through Urban Environments

Hogerbrug, M.; Venrooij, J.; Pool, D.M.; Mulder, Max

**Publication date**
2020

**Document Version**
Final published version

**Citation (APA)**
Hogerbrug, M., Venrooij, J., Pool, D. M., & Mulder, M. (2020). *Simulator Sickness Ratings Reduce with Simulator Motion when Driven Through Urban Environments*. 175-178. Paper presented at Driving Simulation Conference Europe 2020 VR, Antibes, France.

# Simulator Sickness Ratings Reduce with Simulator Motion when Driven Through Urban Environments

**Marc Hogerbrug[1,2], Joost Venrooij[1], Daan M. Pool[2] and Max Mulder[2]**

(1) BMW Group, Research, New Technologies, Innovations, 80807 München, Germany,
e-mail : marc_hogerbrug@outlook.com, Joost.Venrooij@bmw.de
(2) Delft University of Technology, Faculty of Aerospace Engineering, section Control & Simulation, 2629 HS Delft, The Netherlands, e-mail : {d.m.pool, m.mulder}@tudelft.nl

***Abstract -*** *A necessity in driving simulation testing is to understand and attenuate simulator sickness, to reduce the number of undesired drop-outs. Especially urban environments, with its many turns and changes in the velocity profile, are a challenge. This paper describes the motion sickness rating results of a between-subjects experiment ($n = 63$), which investigated the effects of adding scaled yaw motion to a simulator on the sickness incidence and severity while being driven as passenger through an urban environment. Three cases were considered: no motion, scaled yaw motion, and including the vehicle pitch and roll rotations in addition to the scaled yaw motion. The misery scale (MISC) was obtained every minute, and the simulator sickness questionnaire (SSQ) was completed before and after the 45-minute trial. Experimental results show that less participants became sick when some form of yaw-motion was provided.*

***Keywords:*** *Driving simulators, motion cueing, simulator sickness, simulator motion, subjective ratings*

## Introduction

The transition from manually-driven to autonomous cars requires elaborate testing of take-over behavior, cooperative strategies and the use of automation [Hoc18]. Simulator testing provides many advantages over real-life testing, especially in crowded, complex and unpredictable urban environments.

Simulator testing, however, comes at a price: simulator sickness (SS), which is a form of motion sickness (MS). Basically, everyone with a properly functioning vestibular system can be made sick if the provocative stimuli are strong enough [Bir49].

Adding physical motion to a simulator could be a way to reduce simulator sickness, although previous studies show a mix of results, where some positive [Sto11], neutral [Klü15], or even negative [Sha92] effects of simulator motion were reported.

Urban driving is often associated with higher occurrences of sickness. Mourant et al. [Mou07] showed that subjects reported significantly lower MS scores when driving on straight roads in the country and sub-urban environments when compared to driving in city environments. Urban environments are characterized by many (often $90$-degree) turns, high optic flow rates, and larger accelerations.

In this paper we present the effects of adding yaw motion to a simulator, on the SS scores given by subjects who were being driven through an urban environment. Three cases were investigated: no-motion (Case 1), scaled yaw cueing (Case 2), and scaled yaw cueing together with the vehicle pitch and roll motions (Case 3). We applied the misery scale (MISC) [Bos05] and the simulator sickness questionnaire (SSQ) [Ken93].

## Methodology

An experiment has been conducted to investigate the effects of adding scaled, but otherwise veridical, yaw motion to a simulator on simulator sickness. This has been done by performing a between-subjects experiment with $63$ participants, who were driven around in an urban environment as passengers, i.e., they were not driving themselves. The experiment protocol and consent forms were approved by the human research ethics committee of TU Delft.

### Experiment design and simulation

The experiment had just one manipulation, namely the motion provided. Three cases were investigated: the simulator either cued no motion (Case 1), purely yaw motion (Case 2), or roll, pitch and yaw motion (Case 3). Participants were divided in three groups corresponding with these three cases.

Participants were passively driven around an eight-shaped track. An urban environment was simulated including $80$ sharp $90$-degree curves, pedestrians and large city buildings. The scenery can be characterized as eliciting a rather high optic flow.

### Participants

Data were collected from $63$ healthy persons, the characteristics are listed in Table 1. All participants had a driver's license. They were divided such that the age and susceptibility to MS variations were equal among the groups. To achieve the latter, a short version of the motion sickness history questionnaire

(MSHQ) [Gri00] was filled out before the experiment. All participants gave informed consent prior to the experiment. Note that the number of participants in Cases 2 and 3 is slightly higher, as we had to exclude some of the participants in these conditions because of minor technical difficulties, the effects of which could have influenced the measurements.

**Table 1: Participant characteristics.**

| | Case 1: No Motion | | Case 2: Yaw Motion | | Case 3: Yaw, Pitch, Roll motion | |
|---|---|---|---|---|---|---|
| | N | Age | N | Age | N | Age |
| F | 9 | 41.89 (12.33) | 11 | 41 (9.66) | 12 | 39.5 (11.86) |
| M | 11 | 39.09 (12.31) | 10 | 44.5 (11.37) | 10 | 38.6 (10.44) |
| T | 20 | 40.35 (12.40) | 21 | 42.67 (10.66) | 22 | 39.09 (11.24) |

F = female, M = male, T = total

## Apparatus

The driver-in-motion (DIM) simulator of the BMW Group in Munich was used; the main simulator characteristics are listed in Table 2. Subjects were provided with a $220$ degrees outside visual.

**Table 2: DIM simulator characteristics.**

| | Stroke | Max. Velocity | Max. Acceleration |
|---|---|---|---|
| X | 1.08 $[m]$ | 3.7 $[m/s]$ | 37 $[m/s^2]$ |
| Y | 1.00 $[m]$ | 3.2 $[m/s]$ | 35 $[m/s^2]$ |
| Z | 0.22 $[m]$ | 1.6 $[m/s]$ | 35 $[m/s^2]$ |
| Roll | 20 $[°]$ | 135 $[°/s]$ | 2500 $[°/s^2]$ |
| Pitch | 20 $[°]$ | 130 $[°/s]$ | 1000 $[°/s^2]$ |
| Yaw | 45 $[°]$ | 300 $[°/s]$ | 3900 $[°/s^2]$ |

## Measurements

The first measurement was the motion sickness history questionnaire (MSHQ) [Gri00]. For the MSHQ, participants were asked six questions, of which the first five considered the participants' experience with MS while traveling in different forms of transport. Subjects were also asked to give a self-rating of their overall susceptibility to MS compared to other people based on the options below:

1. Clearly less than average,
2. Less than average,
3. Average,
4. More than average, and
5. Clearly more than average.

The MSHQ was only given before the experiment started, to help build three groups of subjects with a similar susceptibility to simulator sickness.

Participants were asked to give a misery scale (MISC) score [Bos05] every minute and also continuously rate the perceived motion incongruence (PMI) [Cle18] during the turns. At the end of the experiment, participants were asked to fill out the Simulator Sickness Questionnaire (SSQ) [Ken93], which served as an extra simulator sickness measure. This paper focuses on the MISC and SSQ scores.

**MISC** The misery scale is an $11$-point Likert-scale which ranges from zero to ten [Bos05], see Table 3. It was used as a verbal rating scale and monitored every minute throughout the experiment trials. Participants were only asked to give a number, which took a couple of seconds, as soon as they were familiar with the rating scale. A copy of the MISC was placed in the car mock-up next to the steering wheel.

**Table 3: Misery Scale (MISC).**

| Symptom | | Score |
|---|---|---|
| No problems | | 0 |
| Uneasiness (no typical symptoms) | | 1 |
| Dizziness, warmth, headache, stomach awareness, sweating, ... | vague | 2 |
| | slight | 3 |
| | fairly | 4 |
| | severe | 5 |
| Nausea | Slight | 6 |
| | fairly | 7 |
| | severe | 8 |
| | (near) retching | 9 |
| Vomiting | | 10 |

**SSQ** The simulator sickness questionnaire (SSQ) by Kennedy, Lane, Berbaum, and Lilienthal [Ken93] can be regarded as the single most important questionnaire regarding simulator, cyber or virtual reality sickness [Kes11]. Its $16$ items were rated on a $4$-point Likert-scale which ranged between "not at all" to "severe". These items are divided over three sub-scales [Ken93] [Kes11]. Eventually, a total SSQ score and scores for every sub-scale can be calculated. The three sub-scales are:

- Nausea (e.g., general discomfort, nausea),
- Oculomotor issues (e.g., eyestrain, focusing issues), and
- Disorientation (e.g., vertigo, concentration issues).

## Procedure

The experiment procedure is summarized in Table 4. Participants first read the instructions, filled out the MSHQ and SSQ questionnaires, and received a briefing before starting the simulations. Two practice drives, one with motion and one without motion, were performed to practice with giving the required ratings. After this familiarization, the Tobii glasses were calibrated and participants were presented with one of the three given cases and driven for a maximum of $45$ minutes. The participants' condition was closely monitored. When the participants indicated a MISC index of $6$ two times in two minutes or a single $7$, the simulation was aborted. After the experiment, the SSQ was filled in again and there was some time for questions of the participants.

## Hypothesis

In this paper we discuss one hypothesis: adding motion to a static driving simulator while being driven in an urban environment will result in less participants becoming sick (H1).

# Results and discussion

Out of the $63$ participants, $5$ participants were excluded, because a too long break was required to restart the system somewhere during the experiment

**Table 4: Experimental procedure.**

| Step | Time [min] | Notes |
|---|---|---|
| Welcome | 5 | |
| Instructions | 3 | Reading |
| Pre-test Interview | 5 | Fill in SSQ/MSHQ |
| Instructions | 2 | Safety instructions |
| Familiarization | 5 | Drive 2 test scenarios |
| Experiment | 45 | Drive 1 Case |
| Post-test Interview | 5 | Fill in SSQ |
| Q&A | 3 | Question and Answers |
| Departure | 2 | |
| Buffer | 15 | For technical failure /other delays |
| Total | 90 | |

execution. After the break, there was not enough time to perform the whole $45$-minute drive again. Because simulator sickness develops over time, this could be a possible confound. Participants that had a break but could still finish the full $45$ minutes without 'jumps' in the data before/after the break were included. For example, a 'jump' in the data could be that the MISC score was $4$ before the break and $1$ after the break. Eventually, $58$ participants were left for the analysis.

## MSHQ ratings

The MSHQ was used to ensure similar MS susceptibility across all groups at baseline. In Fig. 1 a stacked bar plot presents the division of the participants and their indicated MS susceptibility score over the three cases. After excluding five participants that were dropped, it becomes clear that the distribution of participants in Case 3 suffers the most.

In the end, the distribution over the different cases is similar. A Pearson Chi-Square test was performed and resulted in non-significant differences of the MSHQ indications between Cases 1, 2 and 3, $\chi^2(6)$ = $2.67$, $p(0.85) > 0.05$. This outcome substantiates the results that the distribution of the participants over the different cases is fairly done.

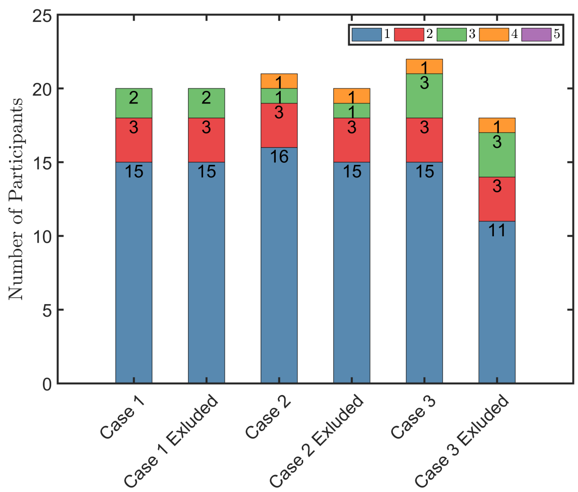


**Figure 1: MS susceptibility division over three cases.**

## Drop-outs

Eight of the 58 participants ($\approx 14\%$) dropped out of the experiment before the end. These participants reported high MISC levels (MISC $>$ 6) at some point in time. Fig. 2 shows that most participants dropped out in the no motion condition (Case 1, $25\%$) and fewer dropped out in Cases 2 and 3. A non-parametric Pearson Chi-square test revealed that these dropout differences were not significantly different, $\chi^2(2)$ = 3.38, $p(0.18) >$ 0.05. When comparing the motion Cases 2 and 3 with the no-motion Case 1, then a nearly significant difference was found, $\chi^2(1)$ = 3.22, $p$(0.073) $>$ 0.025. To prevent a Type 1 error, a Bonferroni correction was applied.

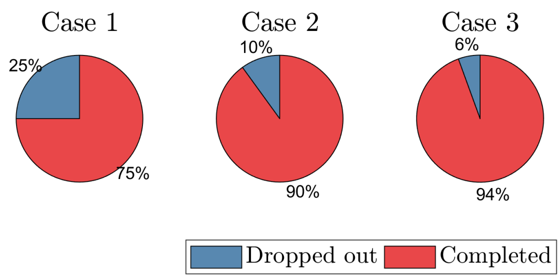


**Figure 2: Number of participants that completed the full experiment or dropped out per case.**

## MISC ratings

Figures 3(a) - (c) show the MISC levels per minute for the three cases. Four groups have been made with different MISC scores in ascending order: 0-1, some uneasiness without clear symptoms; 2-3, some uneasiness with symptoms; 3-5, more severe uneasiness with symptoms; 6-7, participants show nauseated behavior.

Most participants that reach a MISC level of 6 to 7 can be found in Case 1, which is in accordance with the drop-out behavior. Case 3 shows relatively the highest number of MISC levels between 4 and 5, which is already on the edge of getting nauseous. Based on these graphs, Case 2 shows the best results with the majority of participants only reaching MISC levels between 0 and 3. The observed differences between the three cases did not reach statistical significance, however.

## SSQ ratings

The SSQ rating results are shown in Fig. 4; only the total score (TS) is shown.

Generally, the box-plots reveal similar behavior for the 'before' and 'after' experiment ratings of the SSQ. The 'after' SSQ scores' medians are higher for all cases and categories compared to the 'before' SSQ scores median. So, participants became more sick after the simulation trial according to these results. Furthermore, Case 2 shows generally the highest scores for all categories. This is not in line with the MISC ratings which were relatively the lowest compared to the other cases. No significant differences were found between the different cases.

(a) Case 1 (no motion)

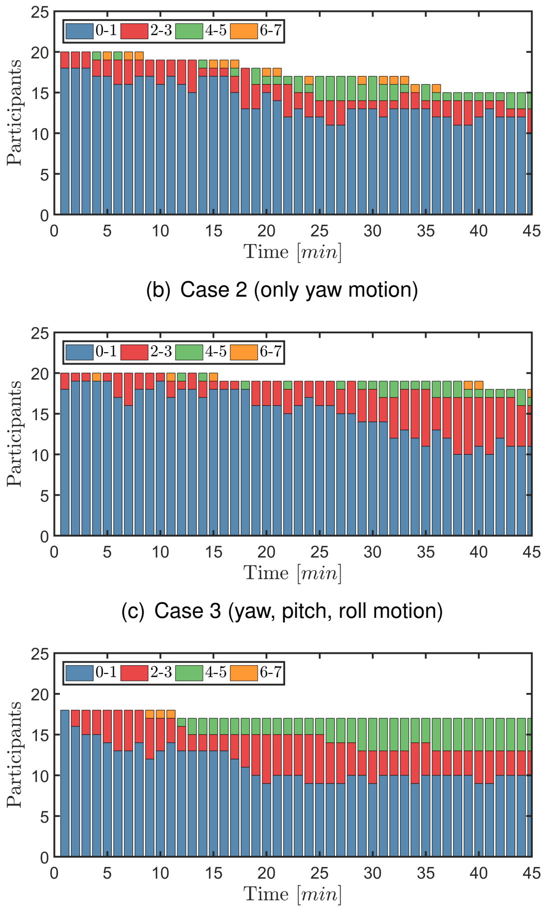


**Figure 3: MISC indications for each minute.**

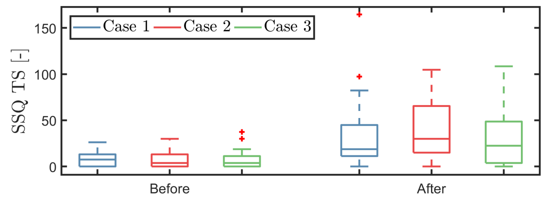


**Figure 4: SSQ Total Score.**

# Discussion

We hypothesized that adding motion to a simulator while driving in an urban environment would result in a decreasing frequency of sick people. When comparing the no-motion case (Case 1) with the motion case (Case 2 and 3) it can be concluded that adding motion to a static driving simulator while driving in an urban environment results in less people becoming sick. Our main hypothesis (H1) is retained.

The MISC scores support H1. More participants reached a MISC score of 6 or 7 for Case 1, which was defined as *sick* in this research, than the other two cases. However, the statistical analysis could not substantiate these results. Collecting more data to boost the proportion of cases falling into each category is a way to resolve this problem.

The SSQ results do not substantiate H1, however, as no difference in SSQ scores was found between the participants of the three cases. The disadvantage of the SSQ is that participants fill it in before and after the stimuli, i.e., when sickness symptoms could have already been partly attenuated. Despite the request to fill in the questionnaire according to the sensations at the end of simulation, it is possible that the subjective judgments of these sensations are inaccurate.

MISC scores are less sensitive to these judgments, since participants are constantly monitored and can be corrected if symptoms arise. In the end, when a participant feels sick, he or she wants to quit the simulation, which is a clear indication of sickness. Therefore, the number of drop-outs and MISC scores are assumed to be more reliable than the SSQ scores.

# Conclusions/implications

We conclude that significantly less participants dropped out in the cases that included simulator yaw motion when driving in urban environments.

# References

J. Birren, **Motion sickness: its psychophysiological aspects. A survey report on human factors in undersea warfare.**, Tech. rep., Committee on Undersea Warfare, National Research Council, Washington, D.C., 1949.

J. E. Bos, S. N. MacKinnon and A. Patterson, **Motion sickness symptoms in a ship motion simulator: effects of inside, outside, and no view.**, *Aviation, space, and environmental medicine*, vol. 76(12): 1111–8, 2005.

D. Cleij, J. Venrooij, P. Pretto, D. M. Pool, M. Mulder and H. H. Bülthoff, **Continuous rating of perceived visual-inertial motion incoherence during driving simulation**, *IEEE Transactions on Human-Machine Systems*, vol. 48(1): 16–18, 2018.

M. J. Griffin and H. V. C. Howard, **Motion Sickness History Questionnaire**, *ISVR Technical Report*, (283), 2000.

P. Hock, J. Kraus, F. Babel, M. Walch, E. Rukzio and M. Baumann, **How to Design Valid Simulator Studies for Investigating User Experience in Automated Driving**, in Proc. of the 10th Int. Conf. on Automotive User Interfaces and Interactive Vehicular Applications - AutomotiveUI '18, 105–117, 2018, ISBN 9781450359467.

R. S. Kennedy, N. E. Lane, K. S. Berbaum and M. G. Lilienthal, **Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness**, *The International Journal of Aviation Psychology*, vol. 3(3): 203–220, 1993.

B. Keshavarz and H. Hecht, **Validating an efficient method to quantify motion sickness**, *Human Factors*, vol. 53(4): 415–426, 2011.

M. Klüver, C. Herrigel, S. Preuß, H.-P. Schöner and H. Hecht, **Comparing the Incidence of Simulator Sickness in Five Different Driving Simulators**, *Driving Simulation Conference 2015*, 16–18, 2015.

R. R. Mourant, P. Rengarajan, D. Cox, Y. Lin and B. K. Jaeger, **The Effect of Driving Environments on Simulator Sickness**, *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 51(18): 1232–1236, 2007.

T. Sharkey and M. E. McCauley, **Does a motion base prevent simulator sickness?**, *Fl. Simulation Techn. Conf.*, 1992.

H. Stoner, D. Fisher and M. Mollenhauer, **Simulator and Scenario Factors Influencing Simulator Sickness**, *Handbook of Driving Simulation for Engineering, Medicine, and Psychology*, 2011.