Effects of Visually Induced Motion Sickness on Emergency Braking Reaction Times in a Driving Simulator

René Reinhard, **Ender Tutulmaz**, **Hans M. Rutrecht**, **Patricia Hengstenberg**, **Britta Geissler**, **Heiko Hecht**, and **Axel Muttray**, Johannes Gutenberg–Universität Mainz, Germany

Objective: The study explores associations of visually induced motion sickness (VIMS) with emergency braking reaction times (RTs) in driving simulator studies. It examines the effects over the progression of multiple simulated drives.

Background: Driving simulator usage has many advantages for RT studies; however, if it induces VIMS, the observed driving behavior might deviate from real-world driving, potentially masking or skewing results. Possible effects of VIMS on RT have long been entertained, but the progression of VIMS across simulated drives has so far not been sufficiently considered.

Method: Twenty-eight adults completed six drives on 2 days in a fixed-base driving simulator. At five points during each drive, pedestrians entered the road, necessitating emergency braking maneuvers. VIMS severity was assessed every minute using the 20-point Fast Motion Sickness Scale. The progression of VIMS was considered in mixed model analyses.

Results: RT predictions were improved by considering VIMS development over time. Here, the relationship of VIMS and RT differed across days and drives. Increases in VIMS symptom severity predicted more prolonged RT after repeated drives on a given day and earlier within each drive.

Conclusion: The assessment of VIMS in RT studies can be beneficial. In this context, VIMS measurements in close temporal proximity to the behaviors under study are promising and offer insights into VIMS and its consequences, which are not readily obtainable through questionnaires.

Application: Driving simulator–based RT studies should consider cumulative effects of VIMS on performance. Measurement and analysis strategies that consider the time-varying nature of VIMS are recommended.

Keywords: driving simulation, simulator sickness, Fast Motion Sickness Scale, generalized mixed models, time-varying covariate

---

Address correspondence to René Reinhard, Fraunhofer ITWM, Fraunhofer-Platz 1, Kaiserslautern, 67663, Germany; e-mail: rene.reinhard@itwm.fraunhofer.de.

HUMAN FACTORS
Vol. XX, No. X, Month XXXX, pp. 1–15
DOI: 10.1177/0018720819829316
Article reuse guidelines: sagepub.com/journals-permissions

INTRODUCTION

Human-in-the-loop driving simulators can present their users with complex driving situations in a controlled and safe manner. This allows for the study of risky driving situations, even when the study's real-world equivalent could put both driver and vehicle in danger, while, at the same time, the complexities of demanding roadway scenarios can be observed (Young, Regan, & Lee, 2008). In addition, these complex scenarios are both standardized and reproducible, with a great amount of control over boundary conditions, potential distractions, and expressed behaviors by other road users, which is not easily achieved in on-road tests (Stoner, Fisher, & Mollenhauer, 2011).

Through these advantages, driving simulators make hazardous and risky driving maneuvers accessible to systematic research while offering a balance between experimental control and ecological validity (Loomis, Blascovich, & Beall, 1999). However, psychological and physiological reactions to the simulator are subject to considerable interindividual differences (Johnson et al., 2011; Reinhard et al., 2017). One individual reaction to driving simulators is the occurrence of induced motion sickness–like symptoms in susceptible users (Kennedy, Lane, Berbaum, & Lilienthal, 1993). This is referred to as simulator sickness or, especially in the context of fixed-base simulators, visually induced motion sickness (VIMS; Keshavarz, Hecht, & Lawson, 2014). Typical symptoms include nausea, difficulty concentrating, or in severe cases, vomiting. Adverse effects of VIMS on driving behaviors have been reported (Helland et al., 2016); however, data regarding effects on braking reaction time have been lacking.

Role of VIMS Progression

Symptoms of motion and simulator sickness are known to worsen with increasing duration of exposure (Keshavarz & Hecht, 2011; Kolasinski, 1995; Lawther & Griffin, 1986; Reinhard et al., 2017) in interindividually different patterns (Bock & Oman, 1982; Davis, Nesbitt, & Nalivaiko, 2015; Keshavarz & Hecht, 2011; Reason & Graybiel, 1970; Reinhard et al., 2017). An earlier evaluation of our data set showed that the extent of VIMS varied depending on the route type and the time of measurement. Both adaptation and habituation were observed (Reinhard et al., 2017) when the validated Fast Motion Sickness Scale (FMS; Keshavarz & Hecht, 2011) was used. This scale enables the assessment of symptom severity in its temporal progression.

Impact of VIMS on Participant Attrition and Driving Behavior

VIMS has been linked to driving behaviors exhibited in simulators, such as steering behavior (Helland et al., 2016) or chosen velocities (Helland et al., 2016; Reinhard, Kleer, & Dreßler, in press). It has been argued that induced symptoms may affect reaction times observed in the simulator (Karl, Berg, Rüger, & Färber, 2013). The influence of VIMS on driving behaviors and experimental results obtained during driving simulations can be affected in several ways (Stoner et al., 2011). One possible path is participant attrition, that is, the loss of participants who withdraw from the study due to the symptoms provoked by the simulator. This can affect particular participant groups disproportionately since factors such as age and gender have been shown to consistently relate to VIMS (Classen, Bewernitz, & Shechtman, 2011). Such groups can also tend toward specific driving behavior (Gwyther & Holland, 2012). The occurrence of sickness-related dropouts can consequently skew experimental results if experimental conditions are not equally provocative with regard to VIMS, thereby excluding participants with a higher probability from the more affected condition. This may affect simulator-based reaction time studies, especially when at-risk populations, such as older drivers, are considered (Edwards, Creaser, Caird, Lamsdale, & Chisholm, 2003). However, even when participants suffering from VIMS do not discontinue the experiment, they may alter their driving behavior in reaction to experienced symptoms. They may, for example, drive slower to alleviate existing symptoms or to avoid further symptom increases (Stoner et al., 2011). Existing symptoms can pose a distraction from the driving task or, through symptoms like concentration difficulties, directly affect task performance (Stoner et al., 2011).

Observed increases in braking reaction times during simulated drives compared to their real-world equivalents have been discussed as possible consequences of induced symptoms (Karl et al., 2013). Given that on-road drivers do not experience equivalent symptoms (Rolnick & Lubow, 1991), their occurrence during simulated drives has been seen as a threat to a driving simulator's absolute validity (Karl et al., 2013; but see Klüver, Herrigel, Heinrich, Schöner, & Hecht, 2016). Here, driving behaviors exhibited in the simulator differ in their absolute observed value from behaviors exhibited during equivalent real-life drives (Mullen, Charlton, Devlin, & Bédard, 2011). Yet even if a simulator's validity with regard to reaction time tests has been established, VIMS could still have a considerable impact on simulator studies.

If VIMS is associated with the behavior under study, it also increases the expressed variance in the driving behavior, which can in turn make it harder to establish effects of interest, even if they actually exist. This is exemplified in a study by Bittner, Gore, and Hooey (1997). Here, emergency braking times observed in a moving-base simulator differed between two visual presentation modalities, but this effect could be shown only when a (nonstandard) simulator sickness measure was included in the analysis as a covariate. This example highlights that the error variance in the statistical test can be decreased by the inclusion of VIMS in the analysis. Consequently, the test's statistical power, that is, its chance of finding an actually existing effect (Cohen, 1988), was increased. Similar relationships between VIMS and reaction times were evident in studies on simple reaction times, obtained after exposure to virtual reality (Nalivaiko, Davis, Blackmore, Vakulin, & Nesbitt, 2015; Nesbitt, Davis, Blackmore, & Nalivaiko, 2017).

As the severity of VIMS can vary considerably across simulated drives, the current study aimed to explore the association of VIMS and braking reaction times over time.

METHODS

Study Design

Twenty-eight participants completed six drives on 2 days in a fixed-base driving simulator. VIMS severity was assessed every minute using the FMS. At five points during each drive, pedestrians entered the road, necessitating emergency braking maneuvers. Mixed models were calculated to evaluate whether VIMS symptom severity improved the prediction of the obtained braking reaction times beyond models that considered only the chosen vehicle speed at the time of the reaction time event and the time spent in the simulator (day, successive driving course, successive reaction time event). The mixed models further explored the role of interaction effects between VIMS and simulator time in the prediction of braking reaction time.

Sample Description

A total of 34 young adults participated in the study and received financial remuneration. Preconditions for participation were (a) an age range from 18 to 30 years; (b) preexisting driving ability, marked by possession of a valid driver's license and a minimum of 5,000 kilometers reported driving experience in right-hand traffic; (c) no indication of previous adaptation to VIMS and no previous simulator experience or extensive computer gaming habits (more than twice weekly, especially racing games); (d) unimpaired eyesight, including normal, or corrected-to-normal, visual acuity and normal color vision; (e) absence of medical histories of heart- or seizure-related conditions; and (f) normal health status, established through medical examinations and general health checkups as well as breath alcohol, urine drug, and saliva caffeine screenings administered by an onsite physician from the Universitätsmedizin Mainz.

Six participants were excluded from the final analysis due to data loss (1 participant), intercurrent illness (1 participant), or marked VIMS (aborted by participant or FMS scores greater than 14; 4 participants). In total, 28 participants completed both days of the experiment (M_{age} = 23.8 years; SD_{age} = 2.5 years; 50% female).

This research complied with the American Psychological Association code of ethics and the tenets of the Declaration of Helsinki and was approved by the local ethics committee. Informed written consent was obtained from each participant.

Apparatus and Materials

Driving simulator. The study utilized a FOERST F10P® fixed-base driving simulator. The setup consisted of a projection (1024 × 788 pixels at 30 Hz; 2000 ANSI-Lumen; dark 3.3 min. Lumen) on a 1.80 × 1.39 m screen (visual angle of projection screen 45.5° horizontal, 35.1° vertical) and a modified Ford Fiesta half-cabin with all standard controls and a built-in audio system. Inside the simulator, temperatures of 22°C and humidity of 45% were maintained through the use of a climate control unit (Hareus Vötsch).

Driving simulation. Participants drove a predefined route on an approximately 23-km-long roadway. First, participants drove 0.6 km on a rural road, followed by driving 0.3 km through a village, during which a cyclist had to be overtaken. Afterward, they entered a 7.1-km-long country road with a car-overtaking maneuver. This was followed by 5.6 km of motorway, which then led into a rural road. The next 3.2 km of rural road contained a segment where participants experienced a rain shower. At the end of the track segment, a bumpy road had to be passed. This was followed by 6.2 km of rural roads interspersed with short drives through village segments. During its latter sections, this part again included a bumpy road as well as hairpin turns.

Lane width and markings were in accordance with German regulations. Generated traffic was present during all road segments, outside of the stretches of road where reaction time events occurred. No pedestrians unrelated to the reaction time events were present. Rural road

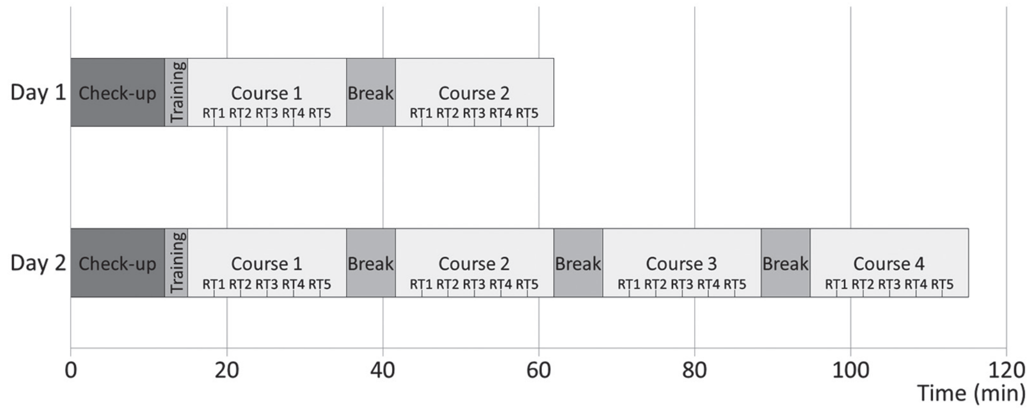

Figure 1. Timeline of the experiment.

segments were surrounded by sporadic trees and bushes, while village drives contained buildings next to the streets. Peripheral objects in the highway section were sparse. On average, completing one driving course took 20.35 min (*SD* = 1.57 min; range = 17–27 min).

Questionnaires. FMS (Keshavarz & Hecht, 2011): A single-item measure of motion sickness symptom severity, ascertained by verbal reports using a numeric response format ranging from 0 (*perfectly fine, no nausea*) to 20 (*extreme nausea, about to vomit*). FMS measurements were prompted by a sound file and a message visible on screen.

Simulator Sickness Questionnaire (SSQ; Kennedy et al., 1993): VIMS symptoms were assessed in greater detail with the SSQ, which consists of 16 individual symptoms, each rated on a four-point response format ranging from *none* to *severe.*

In addition, the participants' subjective fatigue was measured using the Karolinska Sleepiness Scale (KSS; Akerstedt & Gillberg, 1990). At the beginning of both days, the median KSS score was 3 with a 90th percentile of 7.

Procedure

The study comprised two experimental sessions on 2 days separated by at least 1 week and at most 2 weeks. An overview of the experiment's timeline is presented in Figure 1. On each day, the experiment started either at 8:00 a.m. or at 10:00 a.m. with medical tests and checkups. Participants then drove a 3-min-long training session in an environment that was representative of later driving scenarios. They were instructed to adhere to all traffic rules, especially the speed limits. Repeated noncompliance with speed limits led to the experimenter's reminding the participant to observe the traffic rules.

Participants then drove the approximately 23-km-long course. During the drive, participants reported their VIMS symptom severity using the FMS at 20 predefined points, each roughly 1 min apart. The driving course was subdivided accordingly into 19 segments between adjacent FMS measures.

At five points during the drive, a person stepped onto the road from the right side of the street, necessitating an emergency braking maneuver. This person then proceeded to jog straight across the street at a velocity of about 14 km/hr (see Figure 2). At the start of each day, participants were instructed to react to the appearance of pedestrians on the street by performing a braking maneuver. Total braking times, that is, the combination of perception- and motor-related reaction time components (Green, 2000), spanning from the onset of the event until brake application, constituted the dependent variable.

The road-crossing events occurred at prespecified points on the driving course. All reaction time events took place on straight segments of a two-lane rural road with speed limits of 100 km/hr, in accordance with German regulations. The tests occurred under sunny weather conditions on dry roads without unrelated traffic. No other pedestrians were present on the rural road

Figure 2. Reaction time event.

leading to each reaction time–relevant pedestrian crossing event. Before the start of the event, the pedestrian was occluded from the participant's view by a tree, but after he started moving, there were no other obstacles to the pedestrian's visibility. The locations of the five tests were distributed across the drive with each reaction time test falling into separate segments between two FMS measurements, with a minimum of about 2 km between two successive tests. During different drives, each reaction time test was always presented in the same segment between two FMS measurement points. However, to minimize the event's predictability, each individual test location was moved along the same stretch of straight road so that the tests did not occur in the exact same location across drives.

Each event was started dynamically, dependent on the vehicle's current velocity. Here, it was assumed that without action by the driver, the car would have continued on its course at that constant velocity until it hit the pedestrian. The time needed for this, and thus the time the participant had to react to the situation, was calculated as 1.2 s plus the time the vehicle needed to come to a stop from its initial velocity at a negative acceleration of 7.72 m/s^2 after brake application. This time-to-collision had proven to produce an urgent but surmountable task in extensive pretests. On average the reaction time event started at a distance of 65.72 meters (*SD* = 11.40 meters) in front of the car.

On the first day, participants completed two drives in equivalent driving environments that differed slightly only in the placement of the reaction time tests. On the second day, they drove four successive courses according to the same procedure. Between two successive drives, a 5-min break was scheduled. During each break, as well as before the first and after the last drive, participants completed the SSQ and the KSS. After the 5 min had elapsed, participants gave another FMS rating. Depending on the response, the break was extended until a value less than 6 was reported. Of the 112 breaks (28 participants, four breaks that occurred within an experimental day), there were six cases (5.3%) of participants with extended breaks due to increased FMS scores.

Data Analysis Strategy

The data collected from 30 reaction time tests per person (six drives containing 5 reaction time events each) constituted the basis of the following analyses. In total, 795 reaction times were considered in the analysis. That is, 5.4% of trials were excluded from the analysis because the braking response had remained below 50% pedal pressure before the obstacle was passed or due to technical difficulties in the response's recording. The pedal pressure criterion was chosen in accordance with presets by the simulator manufacturer. During training, the participants practiced a strong braking response to pedestrians crossing the road. The pedestrian was detected quickly and reliably by all test persons. Removed trials included eight observed collisions involving pedestrians during the reaction time events, produced by 7 participants.

Data were analyzed using a generalized mixed model approach, which utilized Gamma distributions for reaction time data in compliance with suggestions from the literature (Lo & Andrews, 2015). Gamma distributions were chosen based on both the available literature, showing reaction time distributions to be heavy-tailed continuous unimodal skewed distributions with positive potential values (Green, 2000; Lo & Andrews, 2015) and evaluations of the empirical distributions obtained in this study. Here, the raw reaction time data, as well as their inverse, square root, and natural logarithm transformations, were checked for their deviation from a normal distribution both visually and, supportingly, by using a Shapiro-Wilk test. These distributions were fitted to the raw data using the mass package (Venables & Ripley, 2002) in R, version 3.2.1.

The generalized mixed models were implemented using R's lme4 package (Bates, Maechler, Bolker, & Walker, 2015). The models considered three general sources of influence: (a) the car's velocity at the time when the obstacle entered the street; (b) the reaction time progression over time, manifested in effects of experimental day, driving course, and consecutive reaction time test within each course; and (c) the potential effect of VIMS severity at the time of the test as measured by FMS.

The individual impact of each factor was evaluated by stepwise addition of its terms to the fixed-effect part of the mixed models. This resulted in five models: (a) Model0, the null model, which predicted reaction times using only a group-level intercept; (b) Model1$_{V}$, which accounted only for effects of the velocity at the reaction time event's start; (c) Model2$_{VT}$, which also considered effects of day, course, and reaction time event; (d) Model3$_{VTS}$, where the influence of VIMS severity as measured by the last FMS before the test was added; and (e) Model4$_{VTSi}$, which also considers potential interactions of VIMS with starting velocity and time progression. An appropriate by-subject random effect structure was specified through the principal component analysis–guided evaluation of the uncorrelated random effect terms in the maximal model justified by the design. This was conducted in accordance with the procedure outlined by Bates, Kliegl, Vasishth, and Baayen (2015). The hereby selected by-subject random slopes were related to the effects of successive reaction time event within the drive, the effect of VIMS, and an interaction between the reaction time event number and the experimental day.

The models were compared by evaluating their Akaike information criteria (AIC; Akaike, 1974) and likelihood ratio tests with an α level of .05. In addition, Akaike weights (w[AIC]) and evidence ratios were calculated in accordance with Wagenmakers and Farrell (2004). In the report of individual parameters, bootstrapped 95% confidence intervals are given based on 10,000 samples. The following model comparisons were conducted:

(1) The impact of starting velocity on reaction times was assessed via a model comparison of Model0 and Model1$_{V}$.

(2) A comparison of Model2$_{VT}$ and Model1$_{V}$ investigated the importance of reaction time development over the course of the experiment. This included linear terms relating to day, course, and reaction time event; cross-product terms of the three time-related variables; and their second-order interaction.

(3) Comparisons of Model2$_{VT}$ with Model3$_{VTS}$ and Model4$_{VTSi}$ tested for the predictive value of VIMS-related terms on reaction time predictions beyond the effects of starting velocity and elapsed time. The test against Model3$_{VTS}$ evaluated only the impact of a linear VIMS-related term, while the test against Model4$_{VTSi}$ assessed the combined contribution of both a linear VIMS term and VIMS-related interactions.

Note should be made of the role of vehicle velocity in this comparison. While all participants encountered the same driving situations with the same speed limits, vehicle velocity was still determined by each individual participant. As previously discussed, the chosen velocity has in the past been shown to relate to VIMS (Helland et al., 2016; Reinhard et al., in press). It cannot be ruled out that this relationship arose because participants who were susceptible to motion sickness also preferred to drive at lower driving speeds, even in the absence of reported symptoms. Lower driving speeds in themselves have further been related to the performance in braking reaction time tests (Green, 2000; Jurecki & Stańczyk, 2014; Törnros, 1995). Thus, in the simulator, susceptible users could express VIMS and also choose to drive at lower speeds, which may in turn affect reaction times. The analysis strategy at hand was chosen to preclude these effects.

(4) A comparison of Model4$_{VTSi}$ and Model3$_{VTS}$, which differed only with regard to the cross-product terms of VIMS with starting velocity and experi-

TABLE 1: Median, 10th, and 90th Percentile of Fatigue Reported in Simulator Sickness Questionnaires Obtained After Each Individual Drive

	Day 1		Day 2			
	Course 1	Course 2	Course 1	Course 2	Course 3	Course 4
10th percentile	0	0	0	0	0	0
Median	0	1	0.5	1	1	1
90th percentile	1	1	1	1	2	2

Note. The scale extends from 0 to 3.

ment time variables, was used to further evaluate the contribution of these interaction terms beyond the effects of a linear VIMS-related term.

RESULTS

The mean vehicle velocity at the start of the reaction time event across the experiment was 84.70 km/hr, with 90% of starting speeds falling between 52.69 km/hr and 102.58 km/hr.

The most commonly reported SSQ symptoms over the course of the experiment were headache, nausea, and vertigo. The development of SSQ fatigue over the course of the experiment is reported in Table 1.

The distribution of reaction times was indicated to show significant deviation from the normal distribution both via visual assessment and by a Shapiro-Wilks test, $W = 0.99$, $p < .001$. A Gamma distribution with a shape parameter of 16.78 and a rate of 0.01 was shown to be the most parsimonious approximate descriptive of the reaction time data. The development of mean reaction times over the course of the experiment is displayed in Figure 3a; VIMS experienced at the time of the event is shown in Figure 3b.

Braking Reaction Time Predictions

An overview of all fitted models, including the respective AIC values, can be found in Table 2. The predictions of braking reaction times were significantly improved by considering the velocity at the start of the event ($\Delta AIC_{\text{Model1-Model0}} = -16.01$), $\chi^2(1) = 18.00$, $p < .001$. The model that considered starting velocities was 2995.90 times as likely to be the optimal model among the tested models as the null model. Increases in vehicle velocity were predicted to lead to decreases in reaction times ($\beta_{\text{Velocity|Model1}} = -0.11$; 95% confidence interval [−0.16; −0.06]).

Further addition of variables related to the elapsed time in the experiment, that is, the experimental day, the current consecutive drive on the same day, and the number of the current reaction time event, again improved predictions of reaction times ($\Delta AIC_{\text{Model2-Model1}} = -3.10$), $\chi^2(12) = 29.17$, $p = .004$. Adding time-related variables led to a model that was 4.71 times as likely to be the optimal model.

The prediction of braking reaction times was further improved when VIMS symptoms, in both linear and cross-product terms, were considered ($\Delta AIC_{\text{Model4-Model2}} = -2.07$), $\chi^2(13) = 23.93$, $p = .003$, leading to a model that was 2.82 times more likely to be the optimal model among the set of tested models. By dropping the interaction terms of VIMS with time- and velocity-related terms, the reaction time predictions were worsened disproportionately to the gained simplicity of the model ($\Delta AIC_{\text{Model4-Model3}} = -2.18$), $\chi^2(8) = 18.01$, $p = .021$. In contrast, merely utilizing the linear VIMS term did not improve the predictions of braking reaction times beyond models using only velocity- and time-related variables ($\Delta AIC_{\text{Model3-Model2}} = 0.11$), $\chi^2(5) = 2.35$, $p = .799$. The model including interaction terms was most likely to be the most parsimonious among the tested models, with a likelihood of 56.6%. This was 2.97 times as likely as the model that did not include interaction terms.

The predictions of $\text{Model4}_{\text{VTSi}}$ for an average participant, with mean manifestations of the random slope effects, each time point's mean FMS

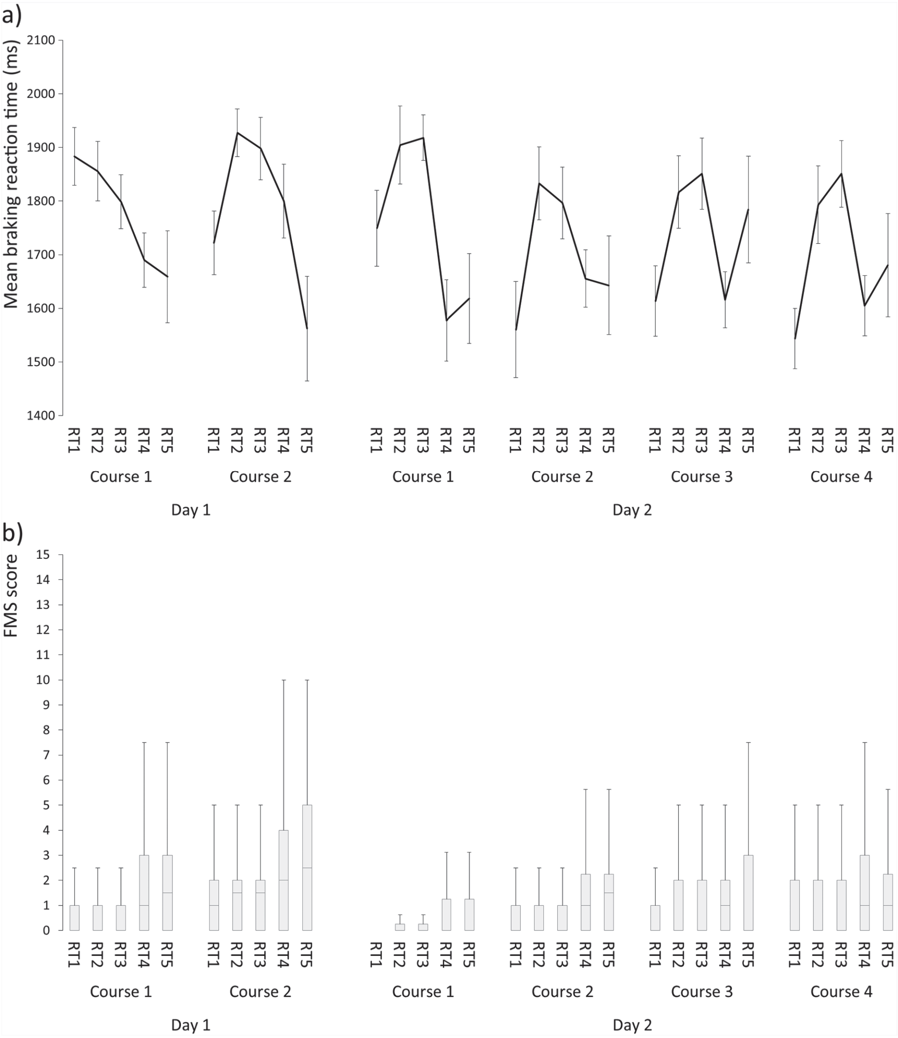

Figure 3. (a) Mean braking reaction times of 28 participants who completed all six courses with five reaction time events per course. The error bars display the standard errors of the mean. (b) Box-Whisker-Plot of Fast Motion Sickness Scale (FMS) scores observed before each reaction time event for 28 participants who completed all six courses. The whiskers extend 1.5 box lengths, unless this exceeds the maximal or minimal observed FMS value.

TABLE 2: The Number of Parameters, Akaike Information Criterion (AIC), and Akaike Weights (w) for All Calculated Mixed Models

Model	Number of Parameters	*AIC*	*w*(AIC)
Model0	3	602.07	.000
Model1$_{V}$	4	586.06	.043
Model2$_{VT}$	16	582.96	.201
Model3$_{VTS}$	21	583.07	.190
Model4$_{VTSi}$	29	580.89	.566

Note. V = velocity; T = time-related; S = visually induced motion sickness (VIMS) symptoms; i = addition of VIMS-related interaction terms.

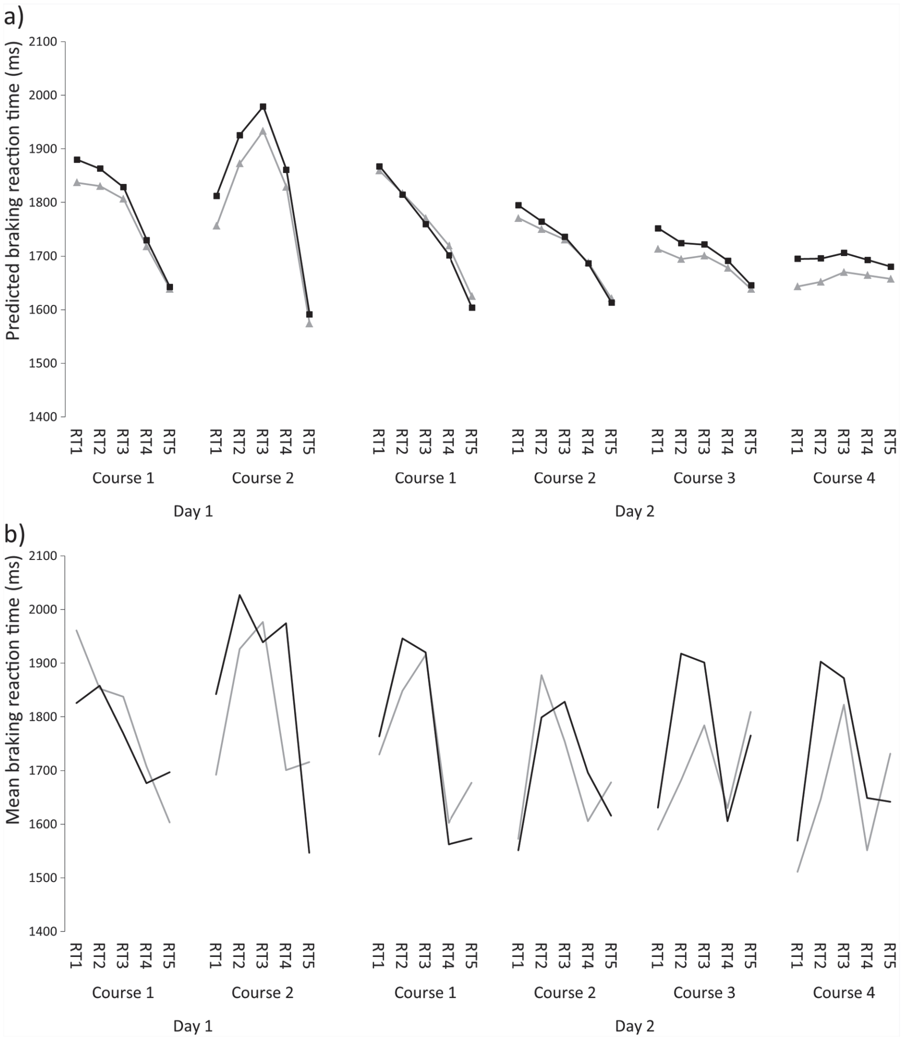

Figure 4. Impact of visually induced motion sickness severity as measured by the Fast Motion Sickness Scale (FMS) on braking reaction times over the course of 30 reaction time events within six driven courses on 2 days. (a) Predicted braking reaction times of a mixed model that considers effects of symptom severity and its interactions with elapsed time in the experiment and starting velocity (Model4$_{\mathrm{VTSi}}$). Grey lines indicate predictions for an average participant, while black lines indicate predictions for a participant who is identical to the average participant except for an increased symptom severity of one point on the FMS at each measurement. (b) Mean braking reaction times for two subgroups with either a higher reported maximum FMS score (FMS$_{\mathrm{max}}$ > 4; black line; $N = 16$) or a lower maximum FMS score (FMS$_{\mathrm{max}}$ ≤ 4; grey line, $N = 12$) according to a median split.

symptom scores, and mean velocities at the start of each braking event, are displayed in Figure 4a. This figure also displays the model's prediction for a participant who is identical to the average participant in every respect except for an increase of the FMS symptom severity by one point at every measurement. As a point of comparison, the observed reaction times for subgroups of participants with higher and lower FMS maxima scores are displayed in Figure 4b.

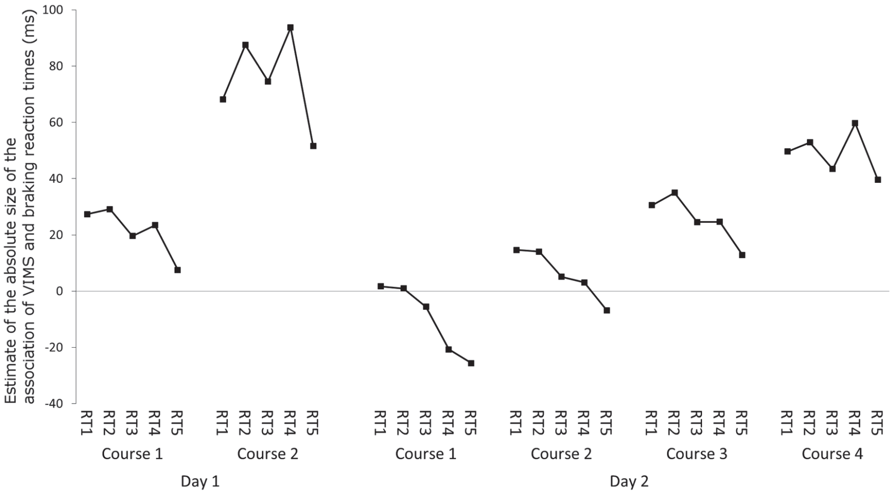

Figure 5. Estimate of the absolute size of the association of visually induced motion sickness (VIMS) and braking reaction times, derived from the product of the mixed model predictions of increases in reaction time per additional VIMS symptom category on the Fast Motion Sickness Scale and the average observed VIMS symptom severity before each of the experiment's reaction time events.

Here, higher FMS maxima describe participants whose maximal FMS score across the experiment was greater than the median of FMS maxima of 4.

$\text{Model4}_{\text{VTSi}}$ predicted that an increase of VIMS by one point led to, on average, prolonged braking reaction times by +20.31 ms. A one-point increase in FMS symptom severity altered reaction time in the range from +50.16 ms to –21.04 ms, depending on when the reaction time event occurred during the experiment. A rough estimate of the association's absolute size can be gained by evoking a comparison standard, which assumes that the observed symptoms are mere artefacts of simulator usage, absent in on-road drives. For this comparison, the average symptom severity observed before each reaction time event was multiplied by the mixed model's prediction for the impact of VIMS for the respective event. The products of the average observed simulator sickness symptoms and $\text{Model4}_{\text{VTSi}}$'s prediction for the impact of each additional point of VIMS symptom severity on reaction times are shown in Figure 5. Among the interactions with symptom severity contained in $\text{Model4}_{\text{VTSi}}$, there was indication that two cross-product terms were different from zero. They related to the interaction of VIMS with the successive reaction time event in the drive ($\beta_{\text{FMSxEvent|Model4}} = -0.02$; 95% confidence interval [–0.02; –0.01]) and with the current drive on each day ($\beta_{\text{FMSxCourse|Model4}} = 0.02$; 95% confidence interval [0.002; 0.03]). For all other interactions, the bootstrapped confidence intervals contained the null value.

DISCUSSION

The study at hand shows that it was useful to consider VIMS symptoms when predicting total emergency braking reaction times in a fixed-base driving simulator. The interaction between VIMS symptoms and time spent in the simulator was important. To analyze this interaction, repeated online VIMS measures were utilized, which assessed experienced symptoms in close temporal proximity to the reaction time events. To make optimal use of this high temporal resolution, generalized mixed models were used in the analysis of observed braking reaction times, specifying a Gamma distribution in accordance with both empirical results and methodological prescriptions (Donmez, Boyle, & Lee, 2008; Lo & Andrews, 2015). It should, however, be noted that despite appreciable deviations from the normal distribution, the shape parameter of the fitted Gamma distribution was high, indicating a high resemblance to a normal distribution (Forbes, Evans, Hastings, & Peacock, 2011).

Association of VIMS With Braking Reaction Times

Mixed model–based comparisons indicate that the consideration of VIMS severity over time improves the predictions of total braking reaction times. This improvement occurs above and beyond any effect that different initial velocities or elapsed simulator time might have. In the current sample, a linear VIMS term did not lead to better reaction time predictions. At the same time, predictions of reaction times were affected by VIMS-related cross-product terms, specifically interactions of symptom severity with the successive drive on each experimental day and with the event within each one of these drives. In the current sample, predicted increases in reaction times due to higher VIMS were more severe during the early road-crossing events of the later drives on each day. This was consistent with the differences between participants with higher and lower FMS maxima according to a median split of the observed VIMS measurements.

For first-time simulator users, the mixed models indicated that the reaction time increase related to one additional point of symptom severity on the FMS was more pronounced (by 15.70 ms) during the second drive compared to the first. When the same participants reentered the simulator a week later, the increases associated with VIMS again differed between the last and the first drive of the day. Here, an additional point of symptom severity led to predictions of prolonged reaction times by an additional 46.19 ms during the last drive. This suggests that even if participants are given the opportunity to rest between drives, events later in the proceedings continue to be affected by VIMS. This was evident despite previous simulator experience and subsequent adaptation (Reinhard et al., 2017).

The differences in individual reaction times linked to the same increase in VIMS also varied between individual reaction time measurements over the course of each drive. When averaged across all instances of the same reaction time event over all drives, an additional point in symptom severity was predicted to lead to average increases in reaction times between 36.88 ms for the first and 3.70 ms for the fifth event. It should be noted that the generally higher symptom levels found during each drive's later events indicate that there still may be a practically relevant absolute difference in reaction times between a drive during which simulator sickness occurs and a drive without symptoms. This is exemplified by the last reaction time event during the first day's second drive (see Figure 5). The finding that a given VIMS symptom increase affects reaction times more during the early events within a drive, as compared to subsequent later events in the same drive, is at first sight surprising. It could merely be a matter of variability. It has previously been indicated that participants may vary strongly in the trajectory of their symptoms over time (Reinhard et al., 2017; see also Bock & Oman, 1982; Davis et al., 2015). Alternatively, the effect could indicate the beginning of a positive adaptation to VIMS. Future studies should therefore explore whether drivers who tend to report early rises in VIMS symptoms differ in their reaction times from those who experience increases in VIMS later during the drive.

Thus, the current study demonstrates potential benefits of assessing VIMS continuously and in close temporal proximity to the behavior under study. The impact that symptom increases have on reaction times does not appear to be uniform over the course of the experiment, meaning that questionnaire approaches, which offer merely a single retrospective measurement for the whole of the drive, may not be optimal in this context (Reinhard et al., in press). Future research could further illuminate how best to obtain FMS measurements close to the reaction time events during simulated drives. Measuring VIMS shortly after a reaction time event would minimize any potential intrusiveness of the measure; however, sharp decelerations, such as emergency braking maneuvers, may lead to increased symptoms (Stoner et al., 2011), and thus the reported symptom severity may overestimate the symptoms experienced during the reaction. If, by contrast, VIMS was assessed shortly before the reaction time event, further research would be needed to see how close to the reaction time event the FMS can be administered without its affecting the driver's reaction during the event.

Simulator sickness has been discussed as a possible impediment to the simulator's validity in reaction time studies (Karl et al., 2013) and as a negative influence on the statistical power of simulator studies (Bittner et al., 1997). A potential impact of symptoms on reaction times obtained in the simulator can impede the simulator's absolute validity, that is, lead to differences in the absolute observed reaction times compared to observations in equivalent real-life drives (Mullen et al., 2011). For the current driving scenario, the mixed models predicted that for the simulated drives observed in this study, with the average observed VIMS symptoms present, the average reaction times tended to be slower compared to predictions for drives without VIMS symptoms. This was consistently observed over the course of the first day, with VIMS being most impactful during the first day's second drive, gaining relevancy again during the second day's later drives. Note, however, that even larger absolute differences between simulated and on-road reaction times in excess of 300 ms have been attributed to VIMS symptoms (Karl et al., 2013). The large differences among studies in this regard, possibly due to differences in simulator and scenario design (Stoner et al., 2011), suggest that future studies should carefully evaluate the expected symptom burden and consider the inclusion of simulator sickness as a dependent variable.

These considerations become especially pertinent with respect to questions of relative validity. To establish relative validity, differences in driving behaviors between conditions of interest must be comparable in both direction and magnitude between real-world and simulated drives (Mullen et al., 2011). There are indications that factors such as sleep deprivation (Kaplan, Ventura, Bakshi, Pierobon, Lackner, & DiZio, 2017) or intoxication (Helland et al., 2016) can influence the level of VIMS expressed in simulators. If these differences in VIMS levels are specific to simulators, they could skew reaction time results obtained in simulator studies between conditions, for example, between sleep-deprived and well-rested drivers. In studies wherein the impact of factors like sleep deprivation on braking reaction times is of interest, simulator sickness should be considered as a potential threat to the study's relative validity and should consequently be carefully monitored.

It should be noted that while the mixed model analyses showed a relationship between simulator sickness symptoms and braking reaction times, further studies are needed to evaluate the causal nature of the association. It is conceivable that VIMS directly affects perception and motor responses but also that the impact on reaction times could, for example, coincide with changes in driving behaviors. One possibility considered in this study was the choice of driving speed. Susceptible participants may tend toward slower driving speeds (Helland et al., 2016; Klüver et al., 2016; Reinhard et al., in press), beyond any effect of VIMS, which could have affected the expressed reaction times (Green, 2000; Jurecki & Stańczyk, 2014; Törnros, 1995). The current analysis strategy was chosen to counteract this possibility through the examination of VIMS effects beyond any effects of the choice of driving speed. However, this could have eliminated potential legitimate components of the VIMS effect from analysis. Here, the indirect impact of VIMS on reaction times through its impact on the choice of driving speed may have been incorporated in the effect of vehicle velocity. The reported model comparisons therefore may have tested only for more direct effects of the expressed symptoms. Future studies could systematically explore the temporal contingencies between vehicle velocity, simulator sickness, and reaction times in greater detail.

The time course of the SSQ subscore for fatigue (see Table 1) showed a small increase at the end of the last two driving courses of the second day. One possible cause is an effect of time on task, which was, however, taken into account in the statistical models. Another possibility is an incipient drowsiness in the context of simulator sickness. The FMS focuses on nausea but not on drowsiness (Lawson, 2014a). Therefore, the possibility that early symptoms of a sopite syndrome (Lawson, 2014b) have been omitted from the models cannot be excluded. Such symptoms also could have an influence on reaction times, parts of which may have covaried with the more nausea-related symptoms measured by the FMS, while other parts may not have been adequately considered. Consequently, the association between FMS scores and reaction times does not necessitate the interpretation of a direct causal effect of nausea on reactive capacity. Further investigations may differentiate effects of early sopite syndrome symptoms from nausea symptoms on reaction time measurements to consider whether the FMS is indeed the optimal measurement instrument in this context.

Finally, the current study analyzed data only from participants who completed all drives, as the number of observed dropouts related to VIMS was insufficient to justify quantitative analyses. Since participant attrition could affect reaction times obtained in driving simulators (Edwards et al., 2003), future studies could also pursue dropout analyses.

CONCLUSION

Driving simulators can serve as a safe and cost-effective research environment in the study of emergency driving maneuvers. Through continuous online measurements of motion sickness, we showed a relationship between VIMS and braking reaction times, which varied over the course of the experiment. The association between VIMS and reaction times was most clearly observed during each day's later drives and for the relatively earlier reaction time events within a given drive. Thus, when designing reaction time studies for driving simulators, VIMS potential should be gauged both as a potential confounding factor and as a detrimental influence on the study's statistical power. Measurements in sufficient temporal proximity to the reaction time tests using online measures such as the FMS, together with an analysis approach that can account for the time-varying nature of VIMS symptoms, can be beneficial.

ACKNOWLEDGMENTS

Parts of the results were obtained in the context of the dissertation project of Ender Tutulmaz.

KEY POINTS

- VIMS was on average associated with prolonged braking reaction times, with the effect varying over the course of the experiment.
- The relationship of VIMS with reaction time was more pronounced during each day's later drives.
- The relationship of VIMS with reaction time was more pronounced during each drive's earlier reaction time events.
- VIMS should be assessed continuously and in close temporal proximity to the behavior under study.
- Appropriate analysis strategies should consider the time-varying nature of VIMS.

ORCID ID

René Reinhard https://orcid.org/0000-0001-9660-7963

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723.

Akerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience*, *52*, 29–37.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, R. H. (2015). Parsimonious mixed models. *arXiv*, *2* , 1–21. Retrieved from https://arxiv.org/abs/1506.04967

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Bittner, A. C., Gore, B. F., & Hooey, B. L. (1997). Meaningful assessments of simulator performance and sickness: Can't have one without the other? *Proceedings of 41st annual meeting of the Human Factors and Ergonomics Society* (pp. 1089–1092). Santa Monica, CA: Human Factors and Ergonomics Society.

Bock, O. L., & Oman, C. M. (1982). Dynamics of subjective discomfort in motion sickness as measured with a magnitude estimation method. *Aviation, Space, and Environmental Medicine*, *53*, 773–777.

Classen, S., Bewernitz, M., & Shechtman, O. (2011). Driving simulator sickness: An evidence-based review of the literature. *American Journal of Occupational Therapy*, *65*(2), 179–188.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Davis, S. L., Nesbitt, K., & Nalivaiko, E. (2015). Comparing the onset of cybersickness using the Oculus Rift and two virtual roller coasters. In Y. Pisan, K. Nesbitt, & K. L. Blackmore (Eds.), *Proceedings of the 11th Australasian Conference on Interactive Entertainment (IE 2015)* (pp. 3–14). Sydney: Australian Computer Society.

Donmez, B., Boyle, L., & Lee, J. D. (2008). Accounting for time dependent covariates in driving simulator studies. *Theoretical Issues in Ergonomics Science*, *9*(3), 189–199.

Edwards, C. J., Creaser, J. I., Caird, J. K., Lamsdale, A. M., & Chisholm, S. L. (2003). Older and younger driver performance at complex intersections: Implications for using perception response time and driving simulation. In M. Rizzo, J. D. Lee, &

D. V. McGehee (Eds.), *Proceedings of the second international driving symposium on human factors in driver assessment, training and vehicle design* (pp. 33–37). Iowa City: University of Iowa.

Forbes, C., Evans, M., Hastings, N., & Peacock, B. (2011). *Statistical distributions* (4th ed.). Hoboken, NJ: John Wiley and Sons.

Green, M. (2000). How long does it take to stop? Methodological analysis of driver perception–brake times. *Transportation Human Factors*, *2*, 195–216.

Gwyther, H., & Holland, C. (2012). The effect of age, gender and attitudes on self-regulation in driving. *Accident Analysis & Prevention*, *45*, 19–28.

Helland, A., Lydersen, S., Lervåg, L.-E., Jenssen, G., Mørland, J., & Slørdal, L. (2016). Driving simulator sickness: Impact on driving performance, influence of blood alcohol concentration, and effect of repeated simulator exposures. *Accident Analysis & Prevention*, *94*, 180–187.

Johnson, M. J., Chahal, T., Stinchcombe, A., Mullen, N., Weaver, B., & Bédard, M. (2011). Physiological responses to simulated and on-road driving. *International Journal of Psychophysiology*, *81*, 203–208.

Jurecki, R. S., & Stańczyk, T. L. (2014). Driver reaction time to lateral entering pedestrian in a simulated crash traffic situation. *Transportation Research Part F: Traffic Psychology and Behaviour*, *27*(1), 22–36.

Kaplan, J., Ventura, J., Bakshi, A., Pierobon, A., Lackner, J. R., & DiZio, P. (2017). The influence of sleep deprivation and oscillating motion on sleepiness, motion sickness, and cognitive and motor performance. *Autonomic Neuroscience: Basic and Clinical*, *202*, 86–96.

Karl, I., Berg, G., Rüger, F., & Färber, B. (2013). Driving behavior and simulator sickness while driving the vehicle in the loop: Validation of longitudinal driving behavior. *IEEE Intelligent Transportation Systems Magazine*, *5*(1), 42–57.

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

Keshavarz, B., & Hecht, H. (2011). Validating an efficient method to quantify motion sickness. *Human Factors: The Journal of the Human Factors and Ergonomic Society*, *53*(4), 415–426.

Keshavarz, B., Hecht, H., & Lawson, B. D. (2014). Visually-induced motion sickness: Causes, characteristics, and countermeasures. In K. S. Kale & K. M. Stanney (Eds.), *Handbook of virtual environments. Design, implementation, and applications* (pp. 647–698). New York, NY: CRC Taylor & Francis.

Klüver, M., Herrigel, C., Heinrich, C., Schöner, H.-P., & Hecht, H. (2016). The behavioral validity of dual-task driving performance in fixed and moving base driving simulators. *Transportation Research Part F: Traffic Psychology and Behaviour*, *37*, 78–96.

Kolasinski, E. M. (1995). *Simulator sickness in virtual environments* (Technical Report 1027). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. Retrieved from www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA295861

Lawson, B. D. (2014a). Motion sickness scaling. In K. S. Hale & K. M. Stanney (Eds.), *Handbook of virtual environments: Design, implementation, and applications* (pp. 601–626). New York, NY: CRC.

Lawson, B. D. (2014b). Motion sickness symptomatology and origins. In K. S. Hale & K. M. Stanney (Eds.), *Handbook of virtual environments: Design, implementation, and applications* (pp. 531–600). New York, NY: CRC.

Lawther, A., & Griffin, M. J. (1986). The motion of a ship at sea and the consequent motion sickness amongst passengers. *Ergonomics*, *29*, 535–552.

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*, 1–16.

Loomis, J., Blascovich, J., & Beall, A. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods, Instruments, & Computers*, *31*, 557–564.

Mullen, N., Charlton, J., Devlin, A., & Bédard, M. (2011). Simulator validity: Behaviors observed on the simulator and on the road. In D. L. Fisher, M. Rizzo, J. Caird, & J. D. Lee (Eds.), *Handbook of driving simulation for engineering, medicine, and psychology* (Chap. 13). Boca Raton, FL: CRC.

Nalivaiko, E., Davis, S. L., Blackmore, K. L., Vakulin, A., & Nesbitt, K. V. (2015). Cybersickness provoked by head-mounted display affects cutaneous vascular tone, heart rate and reaction. *Physiology & Behavior*, *151*, 583–590.

Nesbitt, K. V., Davis, S., Blackmore, K. L., & Nalivaiko, E. (2017). Correlating reaction time and nausea measures with traditional measures of cybersickness. *Displays*, *48*, 1–8.

Reason, J. T., & Graybiel, A. (1970). Progressive adaptation to Coriolis accelerations associated with 1-rpm increments in the velocity of the slow rotation room. *Aerospace Medicine*, *41*, 73–79.

Reinhard, R., Kleer, M., & Dreßler, K. (in press). The impact of individual simulator experiences on usability and driving behavior in a moving base driving simulator. *Transportation Research Part F: Traffic Psychology and Behaviour, SI: Driving Simulation*. Retrieved from https://doi.org/10.1016/j.trf.2018.01.004

Reinhard, R., Rutrecht, H. M., Hengstenberg, P., Tutulmaz, E., Geissler, B., Hecht, H., & Muttray, A. (2017). The best way to assess visually induced motion sickness in a fixed-base driving simulator. *Transportation Research Part F: Traffic Psychology and Behaviour*, *48*, 74–88.

Rolnick, A., & Lubow, R. E. (1991). Why is the driver rarely motion sick? The role of controllability in motion sickness. *Ergonomics*, *34*(7), 867–879.

Stoner, H. A., Fisher, D. L., & Mollenhauer, M. A. (2011). Simulator and scenario factors influencing simulator sickness. In D. L. Fisher, M. Rizzo, J. Caird, & J. D. Lee (Eds.), *Handbook of driving simulation for engineering, medicine, and psychology* (Chap. 14). Boca Raton, FL: CRC.

Törnros, J. (1995). Effect of driving speed on reaction time during motorway driving. *Accident Analysis and Prevention*, *27*(4), 435–442.

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). New York, NY: Springer.

Wagenmakers, E.-J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin & Review*, *11*, 192–196.

Young, K. L., Regan, M. A., & Lee, J. D. (2008). Measuring the effects of driver distraction: Direct driving performance methods and measures. In M. A. Regan, J. D. Lee, & K. L. Young (Eds.), *Driver distraction: Theory, effects, and mitigation* (pp. 85–106). New York, NY: CRC.

René Reinhard received his diploma in psychology from Johannes Gutenberg–University Mainz in 2014. His research interests include the psychological, physiological, and behavioral reactions to driving simulation and virtual reality.

Ender Tutulmaz graduated from the medical school of Johannes Gutenberg–University Mainz in 2016 and specializes in the field of internal medicine. He is currently engaged in his doctoral dissertation in the Department of Occupational, Social, and Environmental Medicine at Johannes Gutenberg–University Mainz.

Hans M. Rutrecht received his BS in psychology from Johannes Gutenberg–University Mainz in 2014. He is currently employed as a crisis communication consultant.

Patricia Hengstenberg, née Nester, graduated from the medical school of Johannes Gutenberg–University Mainz in 2015 and was awarded her doctoral degree from the Institute of Occupational, Social, and Environmental Medicine in 2018. She is currently employed at the Katholisches Klinikum, Mainz, in the field of internal medicine.

Britta Geissler is a specialist in internal and occupational medicine. She received her medical degree in 2009 from the Department of Occupational, Social, and Environmental Medicine at Johannes Gutenberg–University Mainz.

Heiko Hecht holds the chair of experimental psychology at Johannes Gutenberg–University Mainz. He received his PhD at the University of Virginia in 1992 and has has published over 100 journal articles in the domains of motion sickness, virtual reality, artificial gravity, time-to-contact estimation, and intuitive physics.

Axel Muttray is a specialist in occupational and internal medicine, and attained his habilitation in 1999 and adjunct professorship in 2006. He has published numerous articles on topics such as driving simulation, occupational physiology, occupational toxicology, and strains in the work place.

Date received: January 10, 2018
Date accepted: December 19, 2018