Objective measurement of simulator sickness and the role of visual-vestibular conflict situations: a study with vestibular-loss (a-reflexive) subjects

867.5

R.J.V. Bertin, W. Graf, A. Guillot, C. Collet, F. Vienne, S. Espié

LPPA, CNRS/Collège de France, Paris, France
UFR STAPS, Université Lyon 1, France
CIR-MSIS, INRETS, Arcueil, France

Introduction

Driving simulators are being used increasingly for research and development purposes as it is generally safer to evaluate new design or control principles in a mock-up situation. The fact that drivers can experience various and even extreme road conditions in a realistic and risk-free setting makes these simulators appealing also for education, training and even recreation.

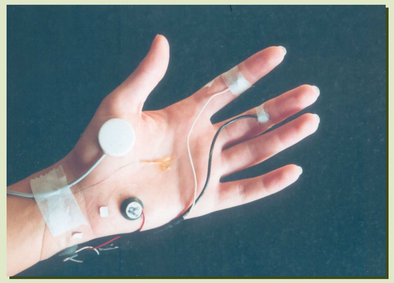
Progress in computer graphics and performance allow for highly realistic simulator *visuals*. High-end models are becoming somewhat better at generating acceptable inertial self-motion information, sometimes even providing real (but limited) linear translation in addition to angular movements. Simpler versions do not generate inertial information at all (fixed-base simulators). A problem that often occurs with all driving simulators is simulator sickness. This phenomenon closely resembles the classically experienced motion sickness and can make a user quit a simulator run within minutes — and cause him/her discomfort for up to several days.

Questions

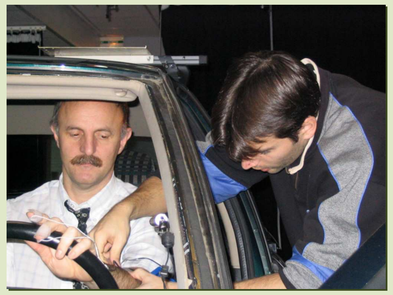
It is commonly accepted that motion sickness is provoked by visual-vestibular conflict — is this true also for simulator sickness? If so, susceptibility in vestibular-loss patients should be significantly less than in controls, as it is for classical motion sickness.

Can the phenomenon be quantified, better understood and ultimately predicted 'on line' through correlation of psychophysical subject reactions and simultaneously recorded neurovegetative activity? Specifically, is it possible to monitor simulator users and detect oncoming sickness before it becomes incapacitating?

I



II



III

IV

Methods

Set-up

We used the fixed-base driving simulator available at the *Institut National de Recherche sur les Transports et leur Securité* (INRETS; Espié, 1999) in Arcueil near Paris (figures I, IV). This simulator provides a large field of view (around 150°) and uses a real car with most controls operative. Linear acceleration is simulated by tilting of the virtual observer (up/down movement in the projected image). It has power steering with force-feedback on the wheel. Driving data (speed and steering commands) were sampled at 30Hz.

Physiological data were recorded using commercially available sensors connected to a recording system developped at Lyon University (Collet *et al.*, 2000). We recorded *skin potential (SP)*, *skin resistance (SR)* and *skin temperature (ST)* using sensors placed on the subjects' left hands and fore-arms (figures II, III). We also recorded instantaneous *heart frequency (HF)* with electrodes placed on the chest. Data were sampled at 10Hz.

The main psychophysical variable recorded was a subjective estimate of discomfort, the *sickness score*. To this end, a visual analog scale was projected continuously in the far low frontal visual field (cf. figure IV). Subjects could indicate their condition by placing a cursor on any of 10 positions between "all is fine" and "I'm about to vomit". The cursor was controlled by command levers on the steering column; it was sampled together with the driving data at 30Hz.

In addition, motion sickness susceptibility was assessed with the Motion Sickness Questionnaire (MSQ) administered before the driving session. Anxiety was assessed three times during the session using Spielberger's STAI, and finally a symptom scoring test (SENSICK) was administered after the experimental session.

Protocol

An experimental session started with the aforementioned MSQ and STAI questionnaires and a brief medical examination to evaluate vestibular and oculomotor function. The driving session was divided into three parts. An initial period of highway driving allowed subjects to familiarise themselves with the simulator, and allow the experimenters to obtain baseline physiolocial activity levels. To minimise the chance of simulator sickness developping, this part was done with only the central of the three screens. After this, the subject took the first of the STAI state anxiety tests, and received instructions for the next driving sequence. This sequence was designed to illicit sickness: to this end, we chose an urban environment with a multitude of sharp corners and traffic lights. It is known that repeatedly taking sharp turns and stopping at a designated point are sickness increasing tasks; in addition, we used the full three screen visual. This sequence consisted of two sessions (*R1, R2*) of approximately 15 minutes separated by a short rest in which the second STAI state test was taken. The third STAI state and the SENSICK test were taken after the second session. The duration was shortened if necessitated by the subject's condition; subjects could terminate the experiment at any time (only 1 subject made use of this). The protocol was approved by the national ethics council (CCPPRB).

Subjects

33 healthy volunteers participated, 23 males aged 24-60y and 10 females aged 25-54y. To date, we have been able to recrute 6 vestibular loss patients, 4 males (29-71y) and 2 females (25 and 29y). All subjects had normal or corrected-to-normal vision, had an active, healthy lifestyle and were (reasonably) experienced drivers.

Data treatment

Driving data (including the sickness score) were resampled to 10Hz and synchronised with the physiological data, with t=0 the start of session part R1. The resulting traces were divided in bins of 10 or 60 seconds and the average sickness score and physiological activity were calculated over these bins. Often, the events of interes, in physiological variables like skin temperature and heart frequency, are narrow but quite strong "peaks". Such signals are easily masked by smaller changes taking place on a much longer timescale (thermoregulation) or by almost equally strong fluctuations on a much shorter timescale. An averaging strategy as just described will tend to filter out these signals. Therefore, we also analysed ST and HF after applying the following filter. Pure integration would convert transient peaks into steps, but a non-zero average activity would cause the integrated signal to grow without bounds. To correct for the average activity (and for feeble long duration changes), we first filtered with a high pass filter with a 60s time constant, before integrating. NB: this is in fact a lowpass filter with identical parameters, and it can thus be applied in real time.

To ensure comparability between subjects, all variables were normalised to the interval [0,1] prior to averaging.

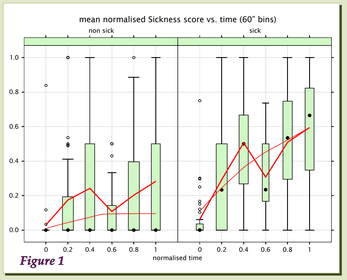


Figure 1

Patients

Out of the 6 patients, one got truly sick (BeGh): he had a high SENSICK score, and also the typical increase in anxiety. Two more subjects developped symptoms (SaRa and CaPo), but were less affected by it and also did not increase their anxiety level. Considering them sick (as in figure 2), there is no significant difference in simulator sickness incidence in the vestibular loss group compared to the control group. If we consider them not sick, there is a marginally significantly lower sickness incidence in the patient group (Kruskal-Wallis, p=0.061).

Results

Qualitative results

As expected, a majority of the participants developped simulator sickness, i.e. 19 out of the 33 control subjects. Here we define "sick" as having requested to terminate or shorten the experiment, or indicating that in real world driving a rest would have been taken. Only one subject terminated the experiment before even starting the R2 session.

Figure 1 shows the normalised sickness score as a function of normalised time in the sick and the non-sick groups; samples at stand-still (v=0m/s) were excluded here. Time was normalised to [0,1] in individual subject recordings to correct for the variable session lengths. The score correlates strongly with this definition of being sick: it is significantly higher in the sick subjects [ANOVA, $F(1,36)=19.6$, $p<1e-4$]. There is a clear, significant (in the sick subjects only) evolution over time [ANOVA, $F(5,169)=13.7$, $p<5e-11$ over the whole population]: the decrease around t=0.6 corresponds to the pause between R1 and R2. Box-and-whiskers displays show median, quartiles and tails; open dots are outliers. Box widths are proportional to the category's N. The thick red lines indicate geometric means and the thin red lines *laess* pairwise linear local fits on the non-categorised observations.

Figure 2 shows the relations between the average sickness score (non-normalised); the increase in anxiety over the duration of R1 and R2; the severity of the symptoms (SENSICK score) and the motion sickness susceptibility (MSQ). Averages for the non-sick (blue) and sick (orange) groups are shown at left with box-whisker plots, per-subject scores are shown in the graph on the right. Significance of the sick/non-sick difference as evaluated with Student's T-Tests are shown in the caption. All tests score higher in the sick population; our definition of sickness corresponds to a SENSICK score of around 8 (out of 80) or higher. Patients are indicated with vertical purple bars.

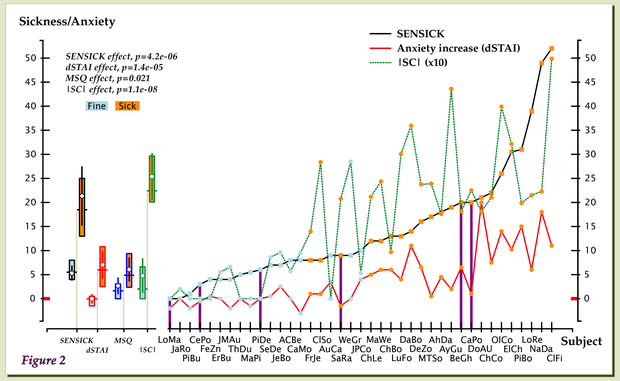


Figure 2

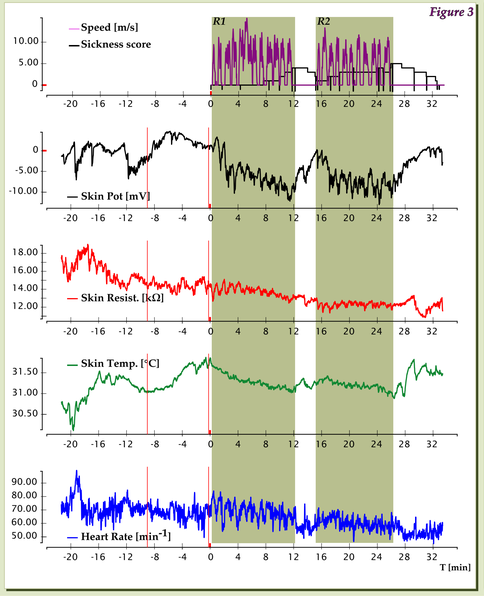


Figure 3

Quantitative results

Figure 3 shows a typical example of combined driving (speed), subjective sickness and physiological data, for subject ChCo. The period t<0 corresponds to the highway driving; the part between the red vertical lines represents baseline activity. Normalisation of the physiological data was performed with respect to the range observed from the leftmost red line (t around -9) to the end of the session. One can see that an increase in the subjective sickness score (top panel, black trace) is accompanied and probably preceded by changes in the physiological parameters. Some, like skin potential return to (almost) baseline during the pause between parts R1 and R2, and after R2. (Vertical lines in the sickness score extending below 0 are synchronisation markers.)

These data were analysed in averaging bins of 10 or 60 seconds; plots of the physiological observables against the normalised sickness score are shown in figure 4. There was no consistent relationship between the sickness score and the skin potential (despite the clear effect visible in figure 3). All parameters tended to decrease with increasing sickness. The presentation of the figures is as in figure 1; the solid blue line shows a linear fit to the non-categorised data (with parameters as indicated); the dashed lines indicate the 95% confidence intervals for slope and intercept. So far, we have analysed the data from 18 subjects.

Figure 4a: the skin resistance is probably the most reliable sickness indicator. It starts decreasing significantly for normalised sickness scores of 0.3 and higher, but remains low once the sickness has 'installed' itself. The main effect of sickness score is highly significant: $F(5,36)=7.3$, $p<1e-4$ (ANOVA).

Figure 4b: the skin temperature also decreases with increasing sickness score, but this is visible only after applying the filter described in the methods. Furthermore, the effect becomes visible only in part R2 (the apparent increase in R1 may reflect thermoregulation and/or the high temperature in the simulator). The decrease in R2 is significant: $F(4,26)=5.95$, $p<0.002$ (ANOVA).

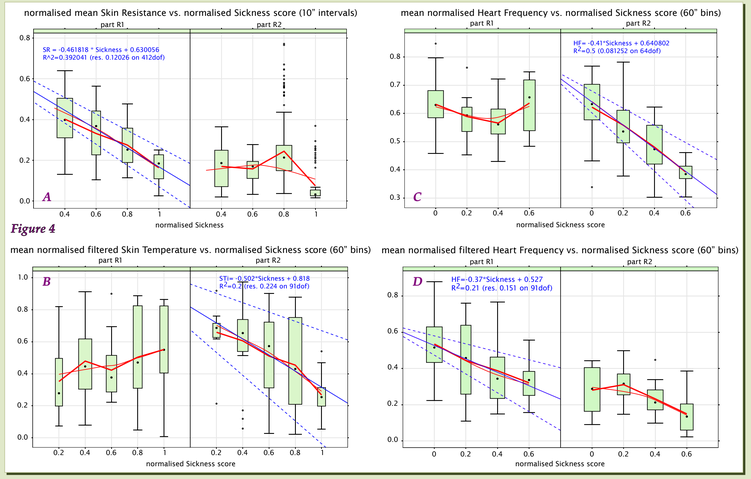


Figure 4

Figure 4c: instantaneous heart frequency also decreases with increasing sickness score, but only up to a normalised score of about 0.7, and only during the R2 part. The effect is significant: $F(3,7)=6.3$, $p=0.02$ (ANOVA).

Figure 4d: after applying the aforementioned filter, there is a significant effect of sickness score on the instantaneous heart frequency in part R1: $F(3,15)=3.47$, $p<0.05$.

Temporal aspects

During the experimental sessions, we had a strong impression that oncoming sickness could be predicted from the evolution of the physiological observables.

Figure 3 supports this impression, but it is not visible from the data analysis as described previously. This is probably due to the fact that subjects reacted on too individual time scales; the R1 and R2 sessions were also of varying lengths. Therefore, we normalised time per session part, such that each part had time in [0,1] seconds. Figure 5 shows skin resistance and temperature plotted versus the per-part normalised time, together with the normised sickness score. There is now a clear indication that the physiological variable decreases before the sharp increase in sickness score begins; for skin temperature, this may even be true in both session parts.

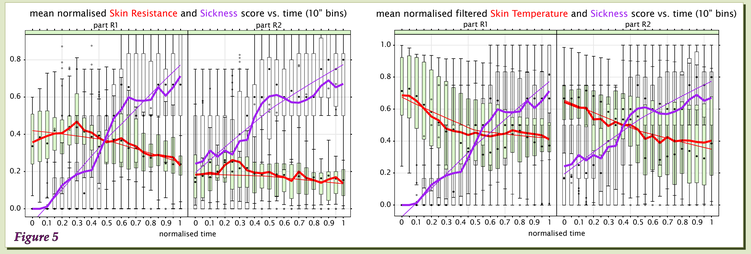


Figure 5

Discussion

General findings

Our study confirmed that simulator sickness is indeed a phenomenon that can pose severe constraints on the applicability of driving simulators. Over 50% of our subjects developped some form of discomfort, and several informed us afterwards that it had lasted several days.

We have shown that this type of sickness is generally accompanied by a significant increase in the state anxiety as assessed by Spielberger's STAI. This test is known to correlate quite well with physiological phenomena.

The simple Motion Sickness Questionnaire can be used to predict simulator sickness susceptibility, but it's use is rather limited: many subjects who never got sick in even the most extreme forms of transport (rollercoasters) did develop simulator sickness.

A number of studies have looked for age and gender effects in motion sickness, and generally found none. In our study, females give lower subjective sickness estimates on average than males (t-test, $p<0.05$). Our data also show an age/gender interaction. Among males, sickness incidence was more or less independent of age, but in females, a strong age effect arose. None of the women under approximately 30 years got sick, whereas all above 30 years did get sick. This effect is highly significant ($p<0.005$), but given the small sample (only 11 females) and the fact that the older age groups are somewhat under-represented, the effect needs to be verified.

Trying to keep a straight course is a known difficult task in simulators because of the lack of accurate seat-of-the-pants inertial information to warn of minute deviations. As a result, many inexperienced users will oscillate, as they try to correct deviations detected visually (= too late). We hypothesised that these oscillations might make the task more nauseogenic, expecting stronger oscillations in the sick subjects. The results seem opposite to that expectation: the power in the steering oscillations tends to decrease with increasing sickness. It is likely that subjects seek to reduce the impact of visual stimulation when getting sick, turning as controlled as possible and avoiding sharp stops and take-offs (which we indeed observed).

Physiological correlates to simulator sickness

There have been earlier attempts at correlating motion or simulator sickness with physiological activity (references below). These studies often failed to find correlation, or found only weak correlation. In contrast, we find strong, reliable correlations between the subjective sickness score and three of the four recorded physiological parameters. In addition, our data seem to support the idea that oncoming sickness can be detected from the autonomous nervous system activity before it becomes unsupportable for the subject. Thus, it would seem to be possible to monitor driving simulator users, and take preventive action when sickness is building up (e.g. imposing a rest, or changing the task). The recent study by Min *et al.* used a paradigm comparable to ours, and found similar results.

Simulator sickness: a visual-vestibular conflict?

In the most optimistic interpretation of our current data, vestibular-loss patients indeed seem less susceptible to simulator sickness than the controls. However, they are clearly not completely insensitive to simulator sickness either. At this point, it can thus be argued that it is not unlikely that a visual-vestibular conflict is involved in the generation of simulator sickness, but it is equally likely that other factors are involved as well. Clearly, we need to expand our patient sample, but already our data show that the phenomenon is an interesting paradigm for the understanding of self-motion perception.

Supported by the European Union (QLK6-CT-2002-00151: EUROKINESIS).

Biaggioni, I., Costa, F., and Kaufmann, H. (1998). *Vestibular influences on autonomic cardiovascular control in humans*. J Vestib Res 8, 35-41.
Bolton, P. S., Wardman, D. L., and Macefield, V. G. (2004). *Absence of short-term vestibular modulation of muscle sympathetic outflow, assessed by brief galvanic vestibular stimulation in awake human subjects*. Exp Brain Res 154, 39-43.
Collet, C., Vernet-Maury, E., Miniconi, P., Chanel, J., and Dittmar, A. (2000). *Autonomic nervous system activity associated with postural disturbances in patients with perilymphatic fistula: sympathetic or vagal origin*. Brain Res Bull 53, 33-43.
Espié, S. (1999). *Vehicle-driven simulator versus traffic-driven simulator: the INRETS approach*. Driving Simulator Conference 99 (DSC'99) proceedings, Paris, France.
Gianaros, P. J., Quigley, K. S., Muth, E. R., Levine, M. E., Vasko, R. C. J., and Stern, R. M. (2003). *Relationship between temporal changes in cardiac parasympathetic activity and motion sickness severity*. Psychophysiology 40, 39-44.
Min, B. C., Chung, S. C., Min, Y. K., and Sakamoto, K. (2004). *Psychophysiological evaluation of simulator sickness evoked by a graphic simulator*. Appl Ergon 35, 549-556.
Mullen, T. J., Berger, R. D., Oman, C. M., and Cohen, R. J. (1998). *Human heart rate variability relation is unchanged during motion sickness*. J Vestib Res 8, 95-105.