Driving Simulators Validation: The Issue of Transferability of Results Acquired on Simulator

Stéphane Espié, Pierre Gauriat, Max Duraz

INRETS-MSIS, 2 avenue du Général Malleret-Joinville
94114 ARCUEIL Cedex – France

Abstract
Simulators have become, and in an inescapable way, a mean to improve the knowledge in the field of driving behaviour. The investigations which can be led with this type of tool concern in particular the driver’s behaviour and the conception of vehicles and road infrastructures.

In the field of driving training the simulator tool is also more and more used, notably for the training (initial and retraining) of professional trucks drivers or special machines drivers (forklifts, cranes, excavators, and also combined harvester, wood feller buncher...). The advantages of the studies on simulators are numerous: lack of risk, reproducibility of the situations, strict control of the parameters, time saving, costs decrease.

The simulated driving situation is nonetheless quite often far from the actual one, due to numerous limits in the simulation of the virtual environment, which limit the ecology of the proposed situation. Furthermore, the lack of risk on a simulator is a major difference between the actual world and the simulated one. The questions of the validity of the results acquired on simulator, and of their transferability towards the actual situations are thus clearly put. These questions are crucial if we want to give credibility to the use of simulators, whether for the study of the behaviour or for training.

The answer is not simple. The tool, which cannot be generic because of its limitations, must be adapted to the use, and validated for this specific use. From this point of view, the new so-called "hi-fi simulator" raises questions, because this denomination is not defined, and is furthermore restrained to a unique identified usage.

The goal of this paper is to review the validation problems and to propose the emergence of "good practices" of usage. We first of all remind the physicals rendering limits of simulators, we then offer a classification of the uses, then we discuss validation problems for each use. We conclude with the necessity of an ethics of use, made more and more necessary by the current and the future possibly wide distribution of the driving simulator tool.
Introduction

Simulators have become an essential means to improve the knowledge in the field of driving behaviour. This kind of tool allows investigations concerning in particular driver’s behaviour, vehicles conception and road infrastructures conception.

In the field of training the simulator tool is more and more used, notably for the initial training and retraining of the professional truck drivers or special machines (forklifts, cranes, excavators but also combine harvester, wood feller buncher...) drivers.

The advantages of the studies on simulators are numerous: absence of risk, reproducibility of the situations, strict control of the parameters, time saving, decrease of costs... The behaviour of the virtual vehicle is nonetheless often very far from the one of the actual vehicle, due to numerous limits in the simulation of the virtual environment and in the rendering devices.

The question of the validity of the results acquired on simulator, and of the transferability towards actual situations, is crucial if we want to assess the credibility of the usage of simulators, whether for behavioural studies or for training.

The answer is complex. The tool, which cannot be generic because of its limitations, must be adapted to the usage, and validated for this specific usage. From this point of view, another question is the recent "hi-fi simulator": this "label" is not defined; moreover it could only be defined for an identified usage.

The aim of our document is to review the problems of validation and to propose the emergence of "good practices" of usage and of “right naming”. We first remind the limits of simulators, we then tempt a classification of the usages, then we discuss problems of validation usage by usage. We conclude on the necessity of an ethics of usage, more and more necessary due to the current (and future) wide distribution of the driving simulator tool.

Limits of the physical rendering

It is for the time being impossible to render with fidelity (i.e. accuracy compared to the actual world) several phenomena (accelerations, brilliant dynamics, etc.). Simulators design thus rests on “tricks” to produce the embedding and the carrying of compromise which guarantees a coherence in the produced sensations.

Accelerations

The mobile platforms have been used for several years in the aeronautical simulators. Gough-Stewart platform with six degrees of freedom is commonly used to simulate the accelerations felt in flight. The technique used to produce the illusion of acceleration is "tilt coordination" (tilt of the platform and thus of the driver for whom the internal sensors are "deceived" and who interprets the gravity as an acceleration in two directions)
[Reid L.D., Nahon M.A., 1985]. This technology, although certified, does not allow a scale 1 rendering and may induce simulator sickness.

The simulation of movement in car is more complex than in aeronautics because of the frequency of occurrence of the phases of acceleration and deceleration, as well as the important dynamics of these phases. An extra complexifying element stems from the contact with the pavement and from the inferred effects (inertials connected to the mass of the vehicle, but also high frequencies connected to the characteristics of the road).

Two very simple examples allow to put in evidence the problems brought up by the rendering in the scale 1 of the accelerations:

The rendering of the braking. A braking up to the stop from a 100 kph speed requires approximately 100m (order of magnitude). The rendering on the scale of the braking would require from the subject a movement of 100m...

The urban situations. It is frequent in urban situation to carry out multiple changes of lane and/or, for example, to carry out a change of lane followed by a 90° bend. The realization of 2 linked changes of lane would require a movement about 7m in side (thus a lateral motion of about 15m), the realization of the change of lane followed by the bend would require much more.

An example of the solutions used for the depiction at scale 1 is the Shinkansen simulator(Japan Railway). This simulator notably uses a side axis of 28m to return the accelerations undergone by passenger in curves.

**Visualisation**

The display in simulator is limited by the graphic engine, by the rendering devices and by the underlying models. These limits concern on one hand the resolution, the luminosity, the colorimetry and the frequency of rendering, and on the other hand the "real time" calculation of the effects of distribution of the light, notably for the meteorological situations "not in clear weather".

The current devices allow a resolution of rendering about 10cpd (cycle per degree) while the visual performance is about 30cpd. The dynamics of projectors luminosity is about 100 cd/m², while the one connected to the sun illumination is roughly 1000000 cd/m² (c.f. figure 1). This point induces a reduction of the contrasts. The spectrum of colours is not covered and thus the shown colours are not faithful. The frequency of resolution, even in 60hz induces jerky movements [Espié, 2001].
Due to these numerous constraints, the visibility on simulator is very limited (we consider that a panel of speed limit is read, by clear time, in 80m on simulator, while approximately 200m in the “actual world”). The overtaking on a bidirectional road in 2 ways is almost impossible to simulate because of the "poor" visibility.

The modelling and the real-time rendering of the static and dynamic lightings is an important problem. The "classic" methods, based on the techniques of ray-tracing, cannot be used in real-time. Solutions exist for lights carried by the subject vehicle (with limits in the case where lights enlighten a enlightened place, an enlightened motorway for instance), and for the simulation of fog situations of [Dumont E. & al., 2004] even with traffic [Cavallo V. & al, 2002]. The question of the lights of the vehicles of the traffic remains.

Usages

Driving simulators are used for different needs. The populations of simulators drivers are also different. We can draw up the following table:

<table>
<thead>
<tr>
<th>Types of drivers</th>
<th>Kind of usage</th>
<th>Vehicle design</th>
<th>Human factors</th>
<th>Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>1- professionals</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.a - With possibility of learning of the simulator</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>1.b - With few possibility of learning of the simulator</td>
<td></td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>2- ordinary people</td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>
We distinguish the professional drivers (car testers, but also professional drivers) from the ordinary drivers, for whom driving is not the main occupation. Among the professionals we distinguish those for whom the simulator is a working tool (car testers) and those for whom it is a tool used for a specific purpose, for example during training courses (professional drivers).

A professional population is generally considered more homogeneous, and the use of simulator can be mandatory. These characteristics explain its identification as a group.

The professionals for whom the simulator is a working tool can learn to use the simulator (and for example to avoid simulator sickness) and to understand the modalities of the transfer of the acquired results towards the reality. This is why they are identified as a group. We must note that for this group the techniques of embedding can be more intrusive (use of virtual reality helmet for example).

**Validation**

Two schools coexist concerning the validation of driving simulators:

1. Intrinsic validation
2. Validation by objective.

We find here globally the designer/user split. The designer tries to demonstrate that his tool is intrinsically valid, the user tries to assess that it is relevant for the usage for which he intends to use it.

**Intrinsic validation**

Here the question is to prove that the simulator is "valid", by comparing the results obtained in simulation and those obtained in actual situation [Reymond G., 2000]. Because of the impossibility evoked earlier to render the totality of the dynamics, one uses mostly transfer functions (essentially a factor of scale).

According to this approach, the simulator is valid if, for example, the accelerations caused by such an operation correspond (in the function of transfer near) to the ones caused by the same operation in the actual world.

The simulator is said "hi-fi" (a designation used by this school) if the error virtual/real is small (but, unlike what exists in sound restitution there is no standard...). Moreover, and as it is the case in sound restitution, all the subsystems have to be “hifi” in order to ensure that the whole system is. Again, for driving simulator, there is no measurement standard.

**Validation by objective**

The question is here to verify the relevance of the tool for a particular usage (study of the driver’s behaviour or training). The object of the study is the human and not the simulator, which is considered here only as a tool. We can study separately the case of the study of the behaviour and the case of the training:

Study of the behaviour - the question here concerns the validity of the tasks carried out by the driver and their transferability towards the reference situation (the actual situation).
The simulator is considered as valid for a given situation if the driving task is "equivalent" to the one carried out in actual situation (at the tactical and operational levels as well as in terms of workload) [Boer E. R., 2000], [Blaauw, G. J., 1982], [Courage C. & al., 2000], [Jamson, H. & Smith, P., 2002]. Unfortunately, a lot of experiments deal only with operational level [Blana E. & Golias J., 1999], [Godley S. T. & al., 2002], [Törnros J., 1998].

The aim is to teach the trainee or to train or retrain a driver. The simulator is considered as valid if the experiences are transferred to the actual driving [Dols J. & al., 2001].

Discussion

It is not possible to reproduce the physics, notably in the visual and accelerations domains. Therefore, every simulator is a specific compromise dedicated to a certain number of usages. It is not possible to consider validation without mentioning the usage and the population of drivers. Moreover, a recurrent problem stems from the technological improvements: each tiny variation of a physical specification of the simulator changes the compromise and can infer (or impose) a modification of the subject's driving strategies (visual cues, dynamic models variations...), thus a re-assessment of the usefulness of the tool.

A professional user of the simulator can learn, during a phase of training which is often long, how to transpose from the virtual situation to the actual situation. He adapts himself to the defects of the simulator he uses, and learns to apply dynamically functions of transfer which allow him to transpose the results into the reality. He also learns to compensate for the sensory incoherence and to avoid simulator sickness. Thus, the simulator can be used for vehicle design: in this case the designer can ignore several elements of the road context as well as the global coherence of the rendering. The simulator is not intrinsically “valid” but can be said so since it is a valuable tool for specific tasks.

An occasional simulator driver has by definition no opportunity to become used to the defects of the tool. The training course or the experiment are limited in time. Thus, the simulator must be conceived in order to privilege a grip in hand as easy as the one of a real vehicle, and to focus more on the relevance of the proposed situation than on the fidelity of the underlying models. The driving situation has to be “accepted” by the drivers and must be as “ecologic” as possible in order to allow transferability. Tasks and workloads have to be as close as possible to the actual ones. Adaptation to the defaults of the tool may induce strategies which differ from the actual one and thus the unusability of results. Experimental psychology proposes tools to address these questions.

Good practices

Good practices for professional use of the simulator is not addressed in this document, since they are specific to each company and since validation is achieved when the simulation results in economical gains.

For training also, the usefulness of the simulators is demonstrated when there is economical gains for the same level of training. The problem here is to ensure that the level of training is equivalent. This point can be only achieved by a scientific comparison
between the classical curriculum and the one using driving simulator. Two groups have to be considered, a reference group (trained with classical curriculum) and a tested one (trained with simulators for all or part of the curriculum). The comparison must take into account not only the short term effects of the training as well as the long term ones (transfer and retention of the acquired knowledge).

For experiments relating to driver behaviours, transferability of the acquired results is the key issue. The problem here is to avoid biased results. Biased results can come either from adaptation to perceptual limitations, or from the use of different strategies (which is in fact also an adaptation). In the first case, the driver will adapt his behaviour while using the same strategies (e.g. he will use a larger time headway). In the second case, he will use a different strategy to achieve his task (e.g. use different clue). If the used clues are different from those used in actual situation, it will be impossible to transfer the acquired results. Unfortunately, that question is seldom addressed in experiments, since it requires a methodology coming from experimental psychology (verbalization conducted by specialist). These biases are difficult to detect but must be, from a scientific standpoint, tracked for each experiment. One can consider that the use of bad clues results from a poor design protocol, as far as database and models are included in the protocol.

Biased results coming from perceptual limitations may induce operational differences between virtual and actual driving [Kemeny A. & Panerai F., 2003]. The limitations are mainly due to rendering devices (and are constraints), but sometimes also from inadequate adjustment (e.g. unsuited driver's eye height): in this case they are independent from the simulators, and depend from the subjects characteristics. To ensure consistent results, protocols have to take into account such points.

**Conclusion**

Driving simulators are nowadays of current use in several applications. Transferability of acquired results is a key issue for credibility of the tool. Because of the numerous uses and of differences within the users, the answer is complex. Each simulator is a dedicated tool, resulting of compromises imposed by cost and technological limitations.

Assessment and transferability depend on usages but have to be checked on a case by case basis. When the use deals with human factors, the tasks and the workload have to be as close as possible to the actual ones. Methodologies coming from the psychological field have to be used to address this point.

Due to reduction of its cost, the tool is more and more widespread. Based on their experience (good and bad), and in order to avoid biased results, « old » users have to promote « good practices » toward the beginners (a kind of framework). The credibility of this message can be ensured only if the community avoids to use « uncontrolled » denomination (i.e. denomination without any scientific definition); the case of the « hifi » denomination being such an example.
One way to achieve this goal would be to try to define a certification for driving simulators and/or for experiments. Unfortunately, this seems to be an utopia, since each simulator is a unique prototype (at least in research institutions) and since experimental design deals with the specificity of those prototypes.

References


