Transportation Research Part F 37 (2016) 78–96

Contents lists available at ScienceDirect

Transportation Research Part F

journal homepage: www.elsevier.com/locate/trf

The behavioral validity of dual-task driving performance in fixed and moving base driving simulators

Malte Klüver [a,c,*], Carolin Herrigel [b], Christian Heinrich [a], Hans-Peter Schöner [a], Heiko Hecht [c]

[a] Daimler AG, Driving Simulators Dept. RD/FFS, Sindelfingen, Germany
[b] MB Tech Group, Sindelfingen, Germany
[c] Department of Experimental Psychology, Johannes Gutenberg-University, Mainz, Germany

ARTICLE INFO

Article history:
Received 29 June 2015
Received in revised form 17 November 2015
Accepted 14 December 2015
Available online 5 January 2016

Keywords:
Driving simulator
Validity
Simulator sickness
Moving-base
Fixed-base

ABSTRACT

Next generation automotive hardware and user interfaces are increasingly pre-tested in driving simulators. What are the potential limitations of such simulations? We determined the relative and absolute validity of five different driving simulators at the Daimler AG by evaluating five functions of an in-vehicle system based on the guideline of the Alliance of Automobile Manufacturers (2006). The simulations were compared to on-road driving. We hypothesized that not only simulator characteristics, but also user characteristics, such as simulator sickness, gender, or age, influence behavioral validity. Even though relating simulator characteristics and user characteristics to driving performance across different driving simulators and driving tasks is difficult, our results are surprisingly in line with the current body of research. We demonstrate the usefulness of all simulators on a relative and partially on an absolute level with moving-base simulators being preferable to fixed-base simulators. As hypothesized, we showed that simulator sickness was significantly associated with impaired performance. In the fixed-base simulators, we found a significant interaction between age and gender, which we could not find in moving-base simulators and in the on-road study. Explanations for our findings and practical implications are discussed.

1. Introduction

Conducting research using driving simulators provides several advantages, compared to real-road studies. Among others, it allows the researcher to (a) be in full control of the driving scenario, (b) examine potentially dangerous situations, (c) repeat an identical experimental condition as often as desired, (d) have access to a vast variety of data, or (e) benefit from large economic savings (Blana, 1996; Classen, Bewernitz, & Shechtman, 2011; Miller & Goodson, 1960).

However, these advantages are only of significance if the results obtained with a driving simulator can be generalized to real-world driving. Even though most validation studies to date demonstrate the usefulness of driving simulators in a variety of research questions (for review see Blana, 1996; Mullen, Charlton, Devlin, & Bédard, 2011), policy-makers still prefer real-road studies (Ranney, 2011).

In this study, for the first time, we directly compare performance in two moving-base simulator setups and three different fixed-base simulators to real-world performance in which participants accomplish secondary tasks while driving. The effect

* Corresponding author at: Daimler AG, HPC X820, 71059 Sindelfingen, Germany. Tel.: +49 151 569 751 87.
E-mail address: malte@kluver.de (M. Klüver).

http://dx.doi.org/10.1016/j.trf.2015.12.005
1369-8478/

of secondary tasks on driving performance or driver distraction received much attention due to its paramount significance for traffic safety (Young & Regan, 2007). For instance, the distracting effects of cell-phone-use (e.g. Caird, Willness, Steel, & Scialfa, 2008; Horrey & Wickens, 2006), operating a radio or cd-player (e.g. Stutts, Reinfurt, Staplin, & Rodgman, 2001), or operating a navigation system (e.g. Chiang, Brooks, & Weir, 2001) on driving performance are well known.

Several attempts have been made to establish standardized evaluation methods of in-vehicle information system (IVIS) safety (for overview see Hurts, Angell, & Perez, 2011), and recently, methodological guidelines have been developed by the Alliance of Automobile Manufacturers (AAM) (2006) and the National Highway Traffic Safety Administration (NHTSA) (2012).

One of the key purposes of this study is to determine the validity of the examined driving simulators. Mostly based on the guidelines of the AAM, we evaluated five functions of an IVIS in an instrumented vehicle as well as in five different driving simulators.

1.1. Types of driving simulator validity

In assessing driving simulator validity, two important distinctions have been established in the literature. First simulator validation approaches distinguished between physical correspondence and behavioral correspondence (Blaauw, 1982), which are nowadays mostly synonymously termed physical and behavioral validity (Mullen et al., 2011). Physical validity refers to the degree to which a simulator reproduces the physical reality. It describes to what extent the physical components such as layout, dynamic characteristics, or visual displays correspond to on-road vehicles. Thus, physical validity focuses solely on simulator characteristics. Physical validity is assumed to be higher in moving-base simulators, compared to fixed-base simulators (Godley, Triggs, & Fildes, 2002; Mullen et al., 2011). Behavioral validity, on the other hand, describes the correspondence between driving behavior in the simulator and on the real road. Thus, behavioral validity refers to the extent to which drivers behave or perform in the same manner as they do on a real road. Most researchers are in agreement about behavioral validity being the more crucial and important type of validity, compared to physical validity (e.g. Blaauw, 1982; Gemou, 2013).

Higher physical validity does not necessarily imply higher behavioral validity. For instance, for lateral motion cueing it was shown that not a scale factor of 100%, but a scale factor between 60% and 70% was preferred when regarding driving performance measures (Pretto, Nusseck, Teufel, & Bülthoff, 2009) or perceived realism (Feenstra, Wentink, Correia Grácio, & Bles, 2009). Additionally, higher fidelity of visual displays has failed to improve driving performance, compared to a lower visual fidelity (Reed & Green, 1999). In other words, there are situations in which a higher physical validity may not improve behavioral validity.

The second widely used distinction in driving simulator validity is between absolute and relative (behavioral) validity (see Blaauw, 1982). Absolute validity is established when dependent variables such as driving parameters, psychophysiological measures, or subjective evaluations take on the same numerical values in a driving simulator as in a real study. Relative validity was originally defined to be established when differences in the dependent variable between conditions are of the same order and direction (Blaauw, 1982). Some authors added to the definition that the magnitude of differences has to be identical as well (Godley et al., 2002; Mullen et al., 2011; Yan, Abdel-Aty, Radwan, Wang, & Chilakapati, 2008). Furthermore, some authors regard relative validity as a qualitative criterion (Blana, 1996), whereas others consider relative validity to be rather a quantitative criterion (Mayhew et al., 2011; Mullen et al., 2011; Wang et al., 2010). However, an explicit and unified definition of relative validity is crucially important for its statistical assessment. Only recently, Wang et al. (2010) addressed this inconsistency by clarifying that when relative validity is defined as having the same order and direction of driver differences, one refers to an identical rank ordering of conditions. However, when relative validity additionally requires the same magnitude of effects to be established, one requires differences between conditions to take on the same numerical values. Thus, it is not necessary to obtain identical numerical values to establish relative validity, but it is crucial for the intervals between conditions to be equal. In this study, we follow the latter definition by regarding relative validity to be established when the differences between conditions are numerically identical.

Whether absolute or relative validity is required depends on the purpose of a study. Since most driving studies follow an experimental design in which the effect of conditions on specific driving parameters is examined, establishing relative validity is usually sufficient (Reed & Green, 1999; Törnros, 1998). However, absolute validity is required when determining absolute numerical values, such as general take-over request times, or when intervention thresholds of an advanced driver assistance system need to be identified (Gemou, 2013). Similar validation studies regarding the safety evaluation of IVIS already exist. Wang et al. (2010) compared performance of three address entry methods in an on-road study to a fixed-base simulator, and concluded absolute validity for visual attention measures and task duration, whereas relative validity was found for lateral and longitudinal vehicle control measures.

Reed and Green (1999) examined the effect of a secondary phone task in an on-road and fixed-base simulator study on driving performance. From their findings, relative validity can be concluded for lane keeping and absolute validity can be concluded for speed variations.

Engström, Johansson, and Östlund (2005) compared performance in a secondary surrogate IVIS task in an on-road study as well as in a fixed-base and a moving-base simulator. They found that in the fixed-base simulator, lane keeping variability was substantially higher, compared to the moving-base simulator and driving on the road, even though results across