# Towards a Common Understanding of Driving Simulator Validity

Chantal Himmels
BMW Group and
Technische Hochschule Ingolstadt
Munich / Ingolstadt, Germany
Johannes Kepler University Linz
Linz, Austria
chantal.himmels@bmw.de

Arben Parduzi
BMW Group
Munich, Germany
arben.parduzi@bmw.de

Martin Fischer
German Aerospace Center
Brunswick, Germany
ma.fischer@dlr.de

Andreas Riener
Technische Hoschschule Ingolstadt
Ingolstadt, Germany
andreas.riener@thi.de

**Figure 1: Example driving simulators with various fidelity. The left two pictures show static driving simulators at the German Aerospace Center in Brunswick (Germany). The middle picture shows the dynamic driving simulator at Technische Hochschule Ingolstadt (Germany). The second picture from the right shows the dynamic driving simulator at the German Aerospace Center in Brunswick (Germany) and the most right picture shows a dynamic driving simulator at BMW Group in Munich (Germany).**

## ABSTRACT

Driving simulators are among the most often used research tools in the AutomotiveUI community. However, no common understanding on when a simulator should be considered valid and how driving simulator validity should be investigated exists, despite numerous publications related to the topic throughout the past four decades. The present paper aims at achieving a more refined understanding of what driving simulator validity actually is. We propose a framework which may be used in context of driving simulator studies and provide recommendations for researchers approaching simulator validation.

## CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction**; *Empirical studies in HCI*; • **Computing methodologies** → Modeling and simulation; Simulation evaluation.

## KEYWORDS

driving simulator, driving simulator validity, framework


*AutomotiveUI '23 Adjunct, September 18–22, 2023, Ingolstadt, Germany*

ACM ISBN 979-8-4007-0112-2/23/09.
https://doi.org/10.1145/3581961.3609895

**ACM Reference Format:**
Chantal Himmels, Arben Parduzi, Martin Fischer, and Andreas Riener. 2023. Towards a Common Understanding of Driving Simulator Validity. In *15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '23 Adjunct), September 18–22, 2023, Ingolstadt, Germany*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3581961.3609895

## 1 INTRODUCTION

Driving simulators hold great potentials to replace real-world testing, given that they are *valid*. As the presented work will highlight, it is still unclear when a simulator should be considered valid and how simulator validity should be investigated. This results in a general lack of understanding of findings on simulator validity and their correct interpretation. Driving simulators are among the most commonly used research tools in the AutoUI community. Simulator studies are useful not only because of the high controllability of the environmental conditions, but also because it is possible to study traffic situations here in an inherently safe test environment [9]. In driving simulators, dangerous driving situations can be explicitly induced and investigated, which would not be possible to examine in real traffic. Nevertheless, a basic understanding of the validity of driving simulators is required in order to understand and assess the opportunities and limitations of both, own and third-party simulator study results. Throughout the present paper, we will discuss different conceptualizations of driving simulator validity which have been proposed in literature. Furthermore, we aim to achieve a better understanding of the influence of the study design on the results obtained. We conclude the paper with recommendations for researchers and practitioners.

## 2 A LITERATURE-BASED DISCUSSION ON DRIVING SIMULATION VALIDITY

Driving simulator validity has occupied researchers since the development of the first simulators. The probably most commonly shared understanding of simulator validity distinguishes between *physical* and *behavioral validity* (e.g., [2]). Physical validity is understood as the physical correspondence of the simulated environment with the real-world vehicle and environment. Behavioral validity refers to the correspondence of driver behavior in the simulator with that in the real world. Behavioral validity is further subdivided into *relative* and *absolute validity* [2]. Absolute validity hereby refers to the numerical correspondence of field and simulator data, while relative validity would be given when effects in the field and in the simulator point in the same direction [2]. Other conceptualizations of simulator validity can be found somewhat less frequently in literature. Blana [4] as well as Engen [16] suggested that the term validity as commonly understood by a psychologist would not match the application in driving simulators. In Psychology, validity is defined as "the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of conclusions drawn from some form of assessment" [32]. In simple terms, this means that validity is given when an experiment methodically permits the conclusions that are drawn on the basis of it. In our opinion, this can also be applied to research in driving simulators, but only if simulator validity is considered in light of the research question and not as a global simulator property. It is generally assumed that a simulator must not always represent all elements of the real environment, but only those that are relevant for it, respectively [9, 15, 21].

The relevance of perceptual correspondence was also frequently mentioned in literature [4, 5], especially in the context of simulator motion [11, 27]. The idea is that perceptual correspondence of (relevant) sensory input cues should result in the same behavior. In presence of perceptual biases, however, driving behavior may be distorted [17]. Internal and external validity [1, 16] refer to the common psychological understanding. Internal validity describes that a construct measures what it is supposed to measure and external validity refers to the degree to which results can be generalized. Ahlstrom et al. [1] further propose statistical validity, referring to the validity of the applied statistical test (also mentioned as statistical conclusion validity by [16]). Validation of realism [1] refers to the extent to which the simulation appears realistic. The latter is similar to what other researchers term face validity [4, 26]. Also construct validity is mentioned in the context of simulator validation [4, 16].

## 3 A FRAMEWORK TOWARDS DRIVING SIMULATOR VALIDITY

Summarizing, there is a wide spectrum of definitions and forms of validity. Note that it is never stated in the listed literature that a specific form of validity is the only one requiring consideration. Anyway, in order to achieve a common understanding, we recommend to rely on the most cited definitions of validity, including physical and behavioral (relative and absolute) validity. We believe that both are rightfully considered, as there are both use cases rather depending on behavioral validity (e.g., most HCI research) and physical validity (e.g., tuning driving dynamics). Taken together, behavioral and physical validity should cover those aspects which are generally relevant (Figure 2).

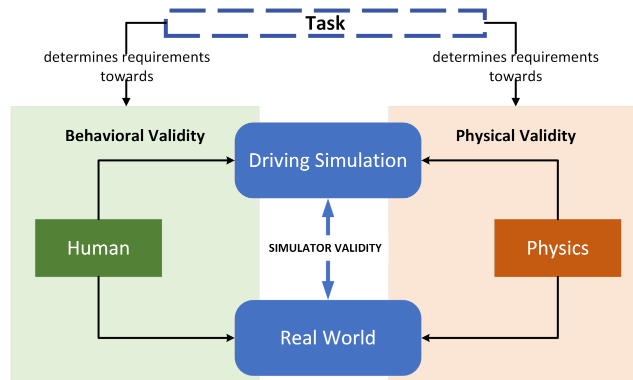


**Figure 2: Suggested driving simulator validity framework, including the aspects behavioral and physical validity (own illustration). Behavioral validity covers those aspects related to the human while physical validity relates to the objective realism of vehicle and environment properties.**

Furthermore, we suggest that validity should always be evaluated considering the use case and research questions [16] based on the fact that not all cues are relevant at all times. We allocate perception as a sub-construct of behavioral validity. It should likewise be influenced by simulator properties and may act as a mediator towards a more realistic driving behavior. Subjective realism, face validity, or the sense of presence may act as a mediator towards behavioral validity as well. Despite increasing subjective realism is a purpose in itself, we do not want to consider it as a form of validity. Note that we further regard the task as the factor defining simulator requirements rather than an influencing factor.

We argue that both perceptual correctness and a minimum sense of presence must be given to obtain valid results, but presume that neither perception nor presence but rather behavioral validity is the key to obtaining valid results. The requirements for a correct statistical methodology are also not specific to simulator validation, but rather a general quality criterion for user studies.

## 4 FACTORS INFLUENCING SIMULATOR VALIDITY

In the following, we will mainly address factors influencing *behavioral validity*, as this aspect is deemed more relevant to the Automotive UI community. The present paper may later be extended to cover more aspects. Extending our framework towards behavioral validity, we suggest **simulator** and **individual** factors to influence behavioral validity and that both aspects should be considered in validation attempts. Simulator properties directly influence physical validity and, via this, behavioral validity (Figure 3).

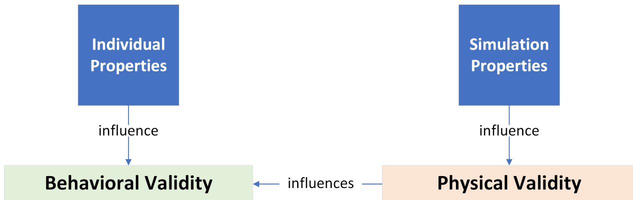


**Figure 3: Simulator properties directly influence physical validity, and, via this, behavioral validity (own illustration). Behavior is further determined by individual factors. Note that even though physical validity is *decreased* by altering a simulator property, behavioral validity may still be *increased*. Simulation properties can hence influence physical validity and behavior via that, yet we do not state that the correlation of physical and behavioral validity must always be positive. For instance, road bumps may be scaled down (decreasing physical validity) to achieve a more realistic driving behavior (increasing behavioral validity) or subjective realism.**

There is no common standard or vocabulary on what elements a driving simulator consists of or what minimum fidelity is required for a setup to be called a driving simulator. A suchlike standard should be agreed on and form the basis for simulator validation studies to foster comparability. Caird and Horrey [9] suggested that the fidelity of a driving simulator and the requirements of a research question should be mapped using a taxonomy or matrix, emphasizing that different research questions come with different requirements.

Published research indicates influences of, for instance, the vehicle mockup [8, 28], the fidelity of the simulated scenario [13], the presence of a motion system [3, 30], motion cueing [7], and the visualization hardware [18, 22] on driving behavior. Note that results differ with regard to different outcome variables, i.e., using a higher fidelity simulator is not generally beneficial but it depends on the research question and the dependent variables considered whether there is an advantage of using a higher fidelity driving simulator.

Also individual influencing factors impact driving behavior in the simulator. Individuals differ, for instance, with regard to their tendency to feel present in virtual environments or their driving style, affecting driving behavior. However, individual influences can hardly be standardized. Issues related to interindividual variance should be overcome using large sample sizes in between-subject designs. In within-subject designs, individual variance is explained variance, though carry-over and contrast effects may occur and require consideration [10, 19].

## 4.1 Mediating factors

As previously noted, **perception** is directly influenced by the physical simulator setup and may influence driving behavior. To date, it is not possible to replicate the full range of perceptual cues which would be present in the real world. Building on the idea that only those cues which are relevant to the task at hand must be replicated, this should also not be necessary. Nevertheless, behavioral validity may be reduced if relevant cues are distorted. Visual depth cues, for instance, cannot be completely adequately represented by displays [6]. To overcome systematic biases, practitioners may decide to intentionally deviate from the correct depiction to achieve more realistic driving behavior [9].

Further, it was hypothesized that the **sense of presence** may act as a mediating factor towards behavioral validity [14, 25]. The sense of presence is defined as the psychological sense of being there in a virtual environment [31]. The basic idea is that the more the virtual driving environment is perceived as being real, the more realistic drivers should behave. While the presented assumption appears rational, note that to the best of our knowledge, there is no research to date proving a positive interrelation of presence and behavioral validity in driving simulation.

## 4.2 Interfering factors

Besides the abovementioned influencing factors, there are further factors interfering with driving simulator validity. The most prominent is probably the occurrence of **simulator sickness** based on the notion that participant behavior changes when experiencing simulator sickness [12]. Simulator sickness is believed to be influenced by individual, task, and simulator factors [23]. Against this context, a simulator inducing a lower incidence of simulator sickness for a specific scenario should be preferred, which is further beneficial to reduce dropouts. Note that [20] did not observe a strong relationship between sickness and driving performance.

Perceptual bias, a lack of a realistic risk perception, or breaks in presence may also negatively impact simulator validity. Mismatches of provided cues, or the relation of different cues, could also impact driving behavior. A 1:1 cueing in longitudinal direction, for instance, may be disadvantageous rather than beneficial when other cues, e.g., the lateral axis is missing or scaled down.

## 4.3 Methodological issues in driving simulator validation studies

Along with the relevance of the use case comes the fact that results can vary with regard to different dependent variables. The operationalization and the selected measure for the construct of interest therefore have a major impact. To put an example, if attention is measured by gaze behavior, different simulator characteristics are likely to play a role (visualisation system, vehicle mockup, ...) than if attention is measured by the standard deviation of the lateral position (steering wheel setup, force feedback, ...).

*Quality of real-world data.* There is always a measurement error in the field [3], which does not necessarily exist in simulation. Second, environmental conditions, such as the behavior of other road users, weather, or lighting conditions, can be perfectly controlled in the driving simulator, whereas being subject to uncontrollable fluctuations in the real world. Third, the place where real-world data are acquired is relevant [22]. If experiments are conducted on test tracks, the transferability to naturalistic driving conditions may be questioned just as much as in the driving simulator, raising the question which test environment is more valid in case of discrepancies. Naturalistic traffic observations constitute an alternative, but offer little controllability.

*Statistical issues in driving simulator validation studies.* Last, the selection of the appropriate statistical test plays an important role. Statistical power must be considered [16]. Validity is most often concluded from the absence of statistical evidence for an effect, i.e., the absence of a significant difference is interpreted as evidence for validity, which is a misconception. We highly recommend that researchers should rely on statistical methods allowing to draw conclusions about *equivalence* instead, such as equivalence tests or Bayesian hypothesis tests [24]. Given that most published research concludes validity from non-significant results of null-hypothesis significance tests, researchers should be aware of the fact that non-significant results may be either due to true equivalence, or a lack of power. Practically speaking, power greatly depends on the sample size, i.e., smaller effects can be detected with a higher probability using larger samples. Also, within-subject designs have a higher power than between-subject designs. We recommend that researchers should first define relevant effects and then decide for an adequate statistical methodology and study procedure.

## 5 RECOMMENDATIONS

Summarizing and in parts extending the information presented above, we conclude our paper with recommendations for both researchers interested in conducting validation studies and researchers and practitioners employing driving simulation as a tool.

**1. What constructs should be considered and measured in context of driving simulator validation?**

To improve the comparability of driving simulator validation studies, a common definition of the constructs related to driving simulator validity should be used. The present work proposes to rely on *behavioral* and *physical validity*.

Beyond that, it is relevant to study both influencing and interfering factors towards behavioral validity, despite we do not consider these as independent forms of validity. Increasing the subjective realism and the sense of presence is a sensible purpose itself, as even if behavior would not be positively affected, compliance and joy of use should increase. In light of ethical considerations, also reducing simulator sickness is a goal detached from simulator validity. Cue mismatches and perceptual correspondence should be investigated considering how these affect driving behavior, as there may be a sweet spot here which does not necessarily go in hand with increasing physical validity.

**2. When should a driving simulator be considered valid?**

First of all, note that a driving simulator itself should not ever be given a label such as "valid" or "invalid", but should only be considered valid or invalid with regard to one (or more, if more have been tested) research question and dependent variables. Installing standards to classify driving simulators, it will become easier in the future to draw inferences from studies conducted by other researchers, as it can be easily evaluated whether the used driving simulators align. The necessity is, however, that researchers make use of and report these standards.

Therefore, when evaluating whether a simulator is valid for a given research question, researchers should consider validation studies focusing on use cases and dependent variables relevant to them rather than "overall" validation attempts. Finally, validity should only be concluded if evidence of statistical equivalence is provided rather than the absence of evidence of effects. Carefully consider sample and effect sizes in studies on simulator validity relying on the results of null-hypothesis significance tests, and evaluate these is light of your own considerations of what effects/differences are practically relevant to your research question. Non-significant results are a lack of evidence, not proof of the absence of effects - and statistical power may simply have been too low to determine relevant differences with null-hypothesis significance tests. With Bayes factors or equivalence tests, consider priors or equivalence bounds in light of those differences which would be practically relevant to your research questions.

**3. How should driving simulator validity be investigated?**

Planning a validation study, it should be considered which scenarios, tasks, and dependent variables are relevant. These first considerations form the basis of the validation study. Regarding statistical methods, we recommend using Bayesian hypothesis tests, as these can provide evidence for both H0 and H1 (within the same test) and discriminate evidence for a null-effect from the absence of evidence. If, despite their shortcomings, null-hypothesis significance tests are used, power analyses should be conducted based on practically relevant effect sizes, and effect sizes should be reported. We also recommend relying on and reporting standards to describe simulator setups as soon as these are available. In this way, a driving simulator can be described objectively.

**4. What driving simulator can be used for a specific use case or research question?**

The question what simulators may be used to validly investigate what research questions is of interest for institutions being able to choose from a selection of simulators, institutions considering which driving simulators to acquire to answer their research questions, and practitioners interpreting the results of driving simulator studies. In the end, researchers will probably be required to create themselves a requirements table for the use cases relevant to them (for illustration, see Table 1). As repeatedly mentioned throughout the present paper, it is recommended to rely on standards for this table, as soon as available. The table can then be filled based on literature and own validation studies. In form of validation studies, simulator properties can be varied systematically (while keeping remaining parameters constant) to achieve an understanding of which simulator properties play a role for the use case under investigation, and to understand how a relevant simulator property should be pronounced.

Practically speaking, we are well aware that we are far from knowing how every simulation property affects each outcome variable in any use case, and that each new combination of simulation properties, which was not explicitly tested, may induce different results. Having conducted a user study for each single simulator property therefore is not a serious prerequisite we can formulate. Also expert knowledge acquired by persons conducting simulator studies as well as literature can be a valuable complement to fill requirement tables.

## 6 LIMITATIONS

This paper is based exclusively on the presented literature. The literature search was not conducted systematically and the existing literature could not be presented in full due to the brevity of the

| | driving dynamics evaluation | user interface testing | safety of use | ... |
|---|---|---|---|---|
| motion base | | | | |
| max. acceleration | | | | |
| image resolution | | | | |
| horizontal field of view | | | | |
| moving objects | | | | |
| steering | | | | |
| ... | | | | |

**Table 1: Example table for a mapping of use cases and requirements. Note that the categories of both simulator properties and use cases are only for illustratory purposes and do not claim correctness or completeness. Depending on the given findings on what simulator properties are required to validly investigate a specific use case (see section 5.4), the table can be filled (e.g., if the use case is "evaluating driving dynamics", the requirements towards a motion base may be quite high, such as six degree of freedom and a maximum acceleration of 6 $m/s^2$, while the requirements towards the visual system may be relatively low, such as 15 pixels per degree resolution, 180° horizontal field of view, ...). These requirements can be assigned to the use cases using a standardized table format.**

publication format. Accordingly, there is no claim to completeness of the discussion presented. The present work also mainly focuses on behavioral validity. It may be extended to cover aspects related to physical validity in the future.

## 7 FUTURE WORK

As repeatedly proposed throughout the present paper, research in context of driving simulation should rely on standards to correctly describe and classify the simulators used. These standards may be applied to match use case requirements and simulator properties. A working group within the German standards association DIN is currently developing a proposal for a suchlike standardized classification of simulators and use cases (a structured table template in the understanding of [9]).

## 8 CONCLUSION

In the present paper, we discussed different definitions of validity and suggested practical recommendations towards conducting validation studies and evaluating the validity of a simulation setup for a question at hand. We recognize that it is unreasonable to demand that every aspect of the impact of the various simulation properties on outcome variables in different use cases be understood prior to conducting a simulator study. Nevertheless, we encourage researchers to carefully question the validity of their given setup before conducting simulator studies and when interpreting the results of simulator studies rather than taking validity as granted.

To conclude our paper, we suggest a rather practical attempt towards driving simulator validity. Sometimes, also the "best", highest-fidelity driving simulator can produce invalid results. Even then it can be valuable to conduct studies in the driving simulator. Selecting a driving simulator, it should also be considered that a higher simulation fidelity drives costs [29]. Efficient planning therefore often means not choosing a higher fidelity than necessary. Subsuming, this paper is intended to initialize further discussions on the topic of simulator validity in order to reach a common understanding and finally standardized methods and criteria for the determination of driving simulator validity.

## ACKNOWLEDGEMENTS

The FLAE approach was applied for the sequence of authors.

## REFERENCES

[1] Christer Ahlström, Anne Bolling, Gunilla Sörensen, Olle Eriksson, and Anders Andersson. 2012. *Validating speed and road surface realism in VTI driving simulator III.* Statens väg-och transportforskningsinstitut.

[2] Gerard J Blaauw. 1982. Driving experience and task demands in simulator and instrumented car: a validation study. *Human Factors* 24, 4 (1982), 473–486.

[3] Evi Blana. 1996. Driving Simulator Validation Studies: A Literature Review. (1996).

[4] Eumorfia Blana. 2001. *The behavioural validation of driving simulators as research tools: a case study based on the Leeds Driving Simulator.* Ph. D. Dissertation. University of Leeds.

[5] Erwin R Boer. 2000. Experiencing the same road twice: A driver centered comparison between simulation and reality. In *DSC2000 (Driving Simulation Conference), Paris, France.*

[6] Jelte E. Bos, Suzanne A.E. Nooij, and Jan L. Souman. 2021. (Im)possibilities of studying carsickness in a driving simulator. In *Proceedings of the Driving Simulation Conference 2021 Europe VR* (2021-09-14), Andras Kemeny, Jean-Rémy Chardonnet, and Florent Colombet (Eds.). Driving Simulation Association, Munich, Germany, 59–63.

[7] Martin Brünger-Koch, Susanne Briest, and Mark Vollrath. 2006. Do you feel the difference?–A motion assessment study. *2006 DSC AP CD-ROM Proceedings* 6, 201 (2006).

[8] GE Burnett, A Irune, and A Mowforth. 2007. Driving simulator sickness and validity: How important is it to use real car cabins. *Advances in transportation studies* 13 (2007), 33–42.

[9] Jeff K Caird and William J Horrey. 2011. Twelve practical and useful questions about driving simulation. *Handbook of driving simulation for engineering, medicine, and psychology* 2 (2011).

[10] Gary Charness, Uri Gneezy, and Michael A Kuhn. 2012. Experimental methods: Between-subject and within-subject design. *Journal of economic behavior & organization* 81, 1 (2012), 1–8.

[11] Diane Cleij, Joost Venrooij, Paolo Pretto, Daan M Pool, Max Mulder, and Heinrich H Bülthoff. 2017. Continuous subjective rating of perceived motion incongruence during driving simulation. *IEEE Transactions on Human-Machine Systems* 48, 1 (2017), 17–29.

[12] Sue VG Cobb, Sarah Nichols, Amanda Ramsey, and John R Wilson. 1999. Virtual reality-induced symptoms and effects (VRISE). *Presence: Teleoperators & Virtual Environments* 8, 2 (1999), 169–186.

[13] Stefan de Groot, Joost CF de Winter, Max Mulder, and Peter A Wieringa. 2011. Nonvestibular motion cueing in a fixed-base driving simulator: Effects on driver braking and cornering performance. *Presence* 20, 2 (2011), 117–142.

[14] Christophe Deniaud, Vincent Honnet, Benoit Jeanne, and Daniel Mestre. 2015. The concept of "presence" as a measure of ecological validity in driving simulators. *Journal of Interaction Science* 3 (2015), 1–13.

[15] C Diels, R Robbins, and N Reed. 2012. Behavioural validation of the TRL driving simulator DigiCar: phase 1: speed choice. In *International Conference on Driver Behaviour and Training, 5th, 2011, Paris, France.*

[16] Thomas Engen. 2008. Use and validation of driving simulators. (2008).

[17] Stéphane Espié, Pierre Gauriat, and Max Duraz. 2005. Driving simulators validation: The issue of transferability of results acquired on simulator. In *Driving Simulation Conference North-America (DSC-NA 2005), Orlondo, FL.*

[18] Yannick Forster, Svenja Paradies, Nikolaus Bee, H Bülthoff, A Kemeny, and P Pretto. 2015. The third dimension: Stereoscopic displaying in a fully immersive driving simulator. In *Proceedings of DSC 2015 Europe Driving Simulation Conference & Exhibition*. Max Planck Institute for Biological Cybernetics Tübingen, Germany, 25–32.

[19] Chantal Himmels, Joost Venrooij, Arben Parduzi, and Andreas Riener. in press. Within- and Between-Subject Designs in Driving Simulator Validation: A Case Study. In *Proceedings of the Driving Simulation Conference 2023 Europe VR*, Andras Kemeny, Jean-Rémy Chardonnet, and Florent Colombet (Eds.). Driving Simulation Association, Antibes, France.

[20] Elizaveta Igoshina, Frank A. Russo, Robert Shewaga, Bruce Haycock, and Behrang Keshavarz. 2022. The relationship between simulator sickness and driving performance in a high-fidelity simulator. *Transportation Research Part F: Traffic Psychology and Behaviour* 89 (2022), 478–487. https://doi.org/10.1016/j.trf.2022.07.015

[21] Nico A Kaptein, Jan Theeuwes, and Richard Van Der Horst. 1996. Driving simulator validity: Some considerations. *Transportation research record* 1550, 1 (1996), 30–36.

[22] Malte Klüver, Carolin Herrigel, Christian Heinrich, Hans-Peter Schöner, and Heiko Hecht. 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 (2016), 78–96.

[23] Eugenia M Kolasinski. 1995. *Simulator sickness in virtual environments*. Vol. 1027. US Army Research Institute for the Behavioral and Social Sciences.

[24] Daniël Lakens, Neil McLatchie, Peder M Isager, Anne M Scheel, and Zoltan Dienes. 2020. Improving inferences about null effects with Bayes factors and equivalence tests. *The Journals of Gerontology: Series B* 75, 1 (2020), 45–57.

[25] Sarah Malone and Roland Brünken. 2021. Hazard perception, presence, and simulation sickness—a comparison of desktop and head-mounted display for driving simulation. *Frontiers in psychology* 12 (2021), 647723.

[26] Robert C McLane and Walter W Wierwille. 1975. The influence of motion and audio cues on driver performance in an automobile simulator. *Human factors* 17, 5 (1975), 488–501.

[27] Gilles Reymond and Andras Kemeny. 2000. Motion cueing in the Renault driving simulator. *Vehicle System Dynamics* 34, 4 (2000), 249–259.

[28] Florence Rosey and Jean-Michel Auberlet. 2014. Driving simulator configuration impacts drivers' behavior and control performance: An example with studies of a rural intersection. *Transportation research part F: traffic psychology and behaviour* 27 (2014), 99–111.

[29] Zwerus Cornelis Roza. 2004. Simulation fidelity theory and practice: A unified approach to defining, specifying and measuring the realism of simulations. (2004).

[30] Shadan Sadeghian Borojeni, Susanne CJ Boll, Wilko Heuten, Heinrich H Bülthoff, and Lewis Chuang. 2018. Feel the movement: Real motion influences responses to take-over requests in highly automated vehicles. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–13.

[31] Mel Slater and Sylvia Wilbur. 1997. A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments* 6, 6 (1997), 603–616.

[32] Gary R VandenBos. 2007. *APA dictionary of psychology*. American Psychological Association.