DSC 2024 EUROPE VR
Driving Simulation & Virtual Reality Conference & Exhibition
18-20 September 2024
Palais des Congrès et de la Musique, Strasbourg | France

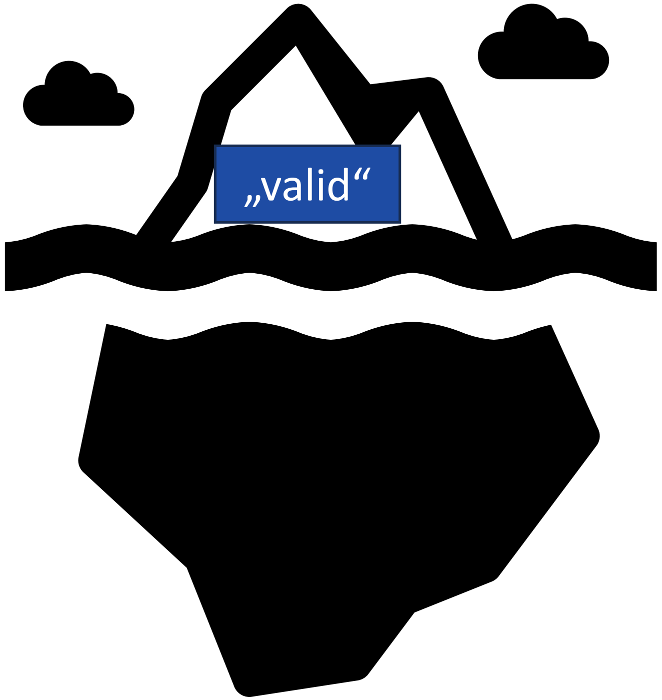
Practical Considerations Towards a Common Understanding of Driving Simulator Validity

Chantal Himmels, Martin Fischer, Arben Parduzi, and Andreas Riener

Deutsches Zentrum für Luft- und Raumfahrt
German Aerospace Center
Technische Hochschule Ingolstadt
JOHANNES KEPLER UNIVERSITY LINZ

Predecessor Paper: AutoUI ´24

Himmels, C., Parduzi, A., Fischer, M., & Riener, A. (2023, September). Towards a Common Understanding of Driving Simulator Validity. In *Adjunct Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 191-196).



under the surface

ambiguity regarding influencing factors

quality and availability of real-world data

unclear definition of term „validity“

inappropriate statistical methods

use case dependence

Predecessor Paper: AutoUI ´24

Himmels, C., Parduzi, A., Fischer, M., & Riener, A. (2023, September). Towards a Common Understanding of Driving Simulator Validity. In *Adjunct Proceedings of the 15th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 191-196).

RECOMMENDATIONS

1 DEFINITIONS & CONSTRUCTS

- Use **behavioral** and **physical validity** to allow for comparability across different studies and papers
- Increasing the **sense of presence** and **behavioral realism** is self-evident
- Reducing **simulator sickness** is relevant for ethical reasons and to prevent dropouts

2 WHEN IS A SIMULATOR VALID?

- Validity should only be considered in light of the **use case**
- In line, validation studies must also attempt realistic use cases
- Rely on and report **standards** as soon as available
- Rely on statistical equivalence, practically considering relevant effects. **Do not conclude validity from non-significant results**

3 INVESTIGATING VALIDITY

- determine **relevant use cases**
- **Bayesian hypothesis tests** are helpful
- When using null-hypothesis significance tests, consider statistical **power** and report **effect sizes**

4 SELECTING A SIMULATOR

- Based on **studies and literature**
- Rely on **standards**
- **Expert knowledge**

Lack of standards for driving simulators

NASA-STD-7009

- **Validation:** degree to which the simulation is an accurate representation of the real world
- **Verification:** degree to which a computational model represents a solution for the intended use
- **Credibility:** confidence in simulation results
- **Uncertainty:** possible deviation of the simulation from the true value

ICAO Manual of Criteria for the Qualification of Flight Simulation Training Devices

→ a standard on what simulators are required to simulate what use cases.

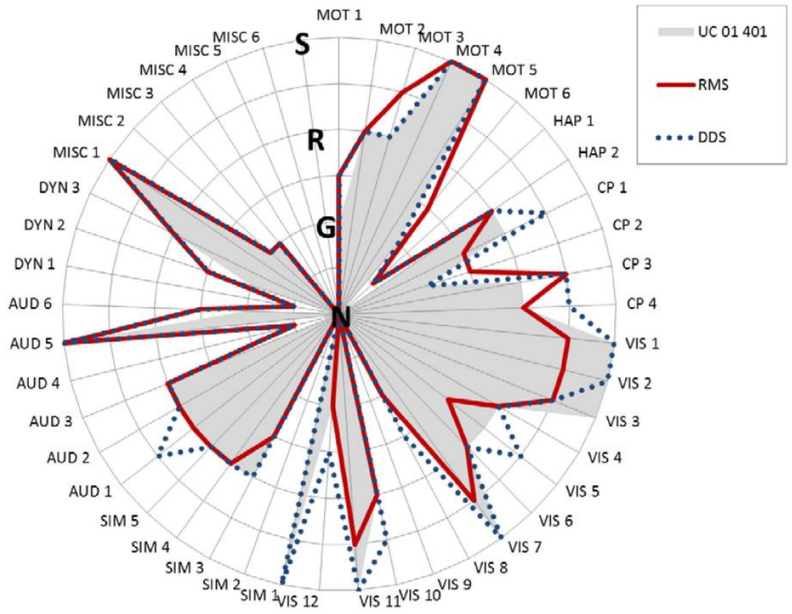
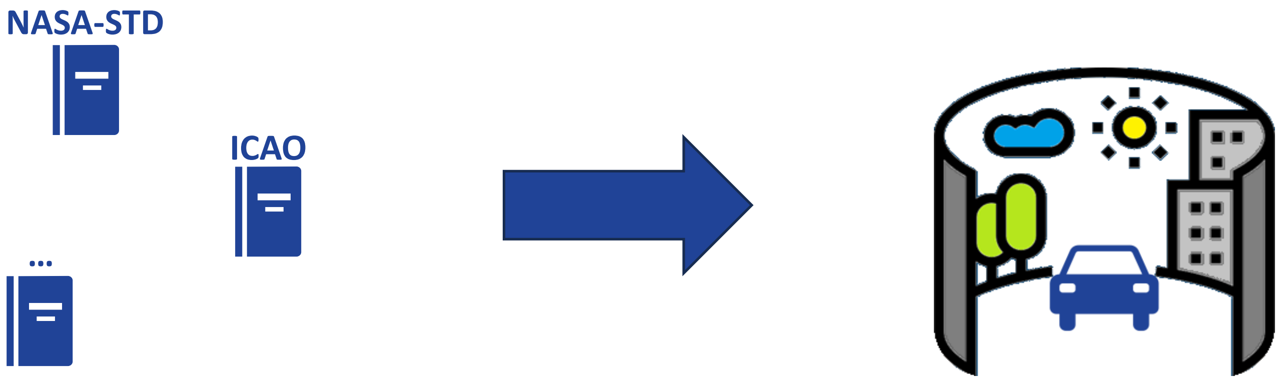


Figure 11: Requirement specification and simulator evaluation regarding the ADAS testing scenario

Fischer, M., Labusch, A., Bellmann, T., & Seehof, C. (2015, September). A task-oriented catalogue of criteria for driving simulator evaluation. In *Proceedings of the Driving Simulation Conference 2015* (pp. 139-150).

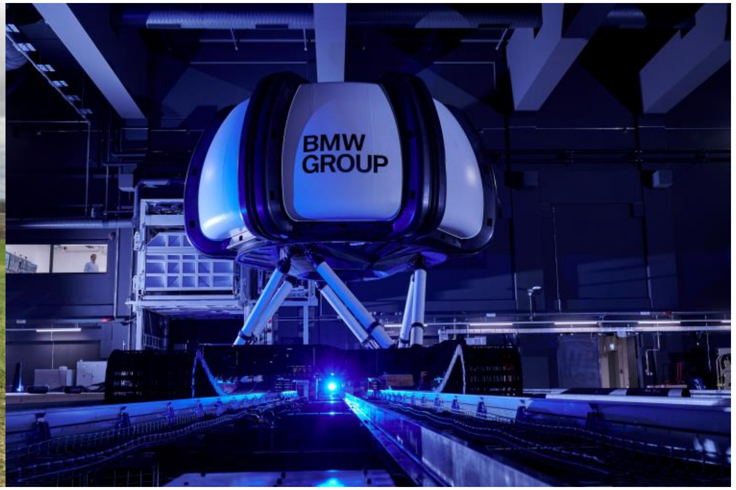
Transfer of standards and best practices to driving simulator studies

- Currently, driving simulator studies rarely refer to or make use of the constructs defined for the aviation and space simulation domains, even if benefits may be expected
- Within the scope of the present paper, we aim to demonstrate how a suchlike transferal of learnings can be expected.



Example I: Validating acceptance in context of level 2 driving automation

- A validation study was performed at BMW Group in Munich, utilizing our high-fidelity driving simulator and a real-world test vehicle
- Investigated acceptance, trust, mental model, … in context of a country road drive



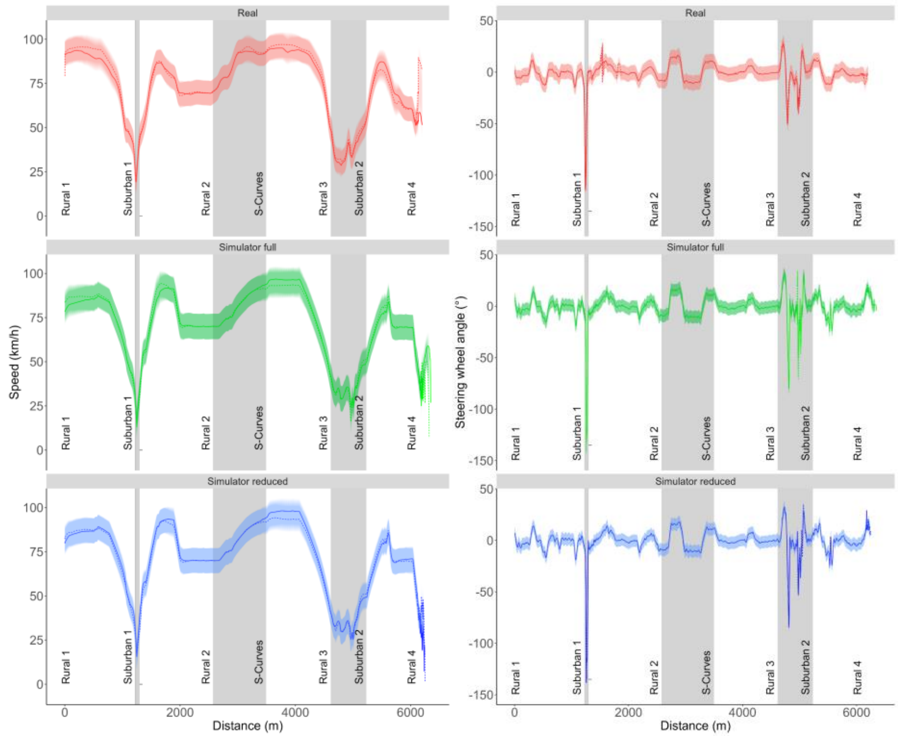
Example I: Validating acceptance in context of level 2 driving automation

VALIDATION

- Real-world comparison study
- Digital twins
- Bayesian methods to determine statistical equivalence

VERIFICATION

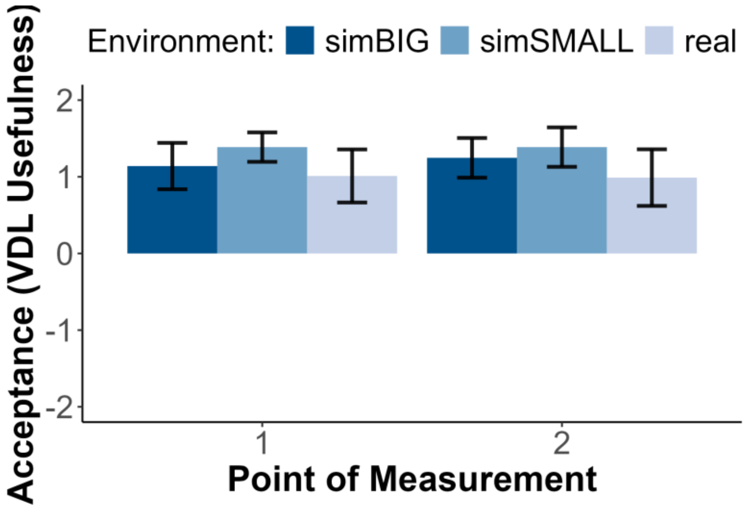
- Verification of dynamics model accuracy and comparison of automated driving controller



Example I: Validating acceptance in context of level 2 driving automation

CREDIBILITY

- The produced outcomes may be trusted with regard to the given setup
- Example:



UNCERTAINTY

- The given findings point at a sensitivity of the results to the fidelity of the **motion system**.
- We are hence uncertain about the credibility of results achieved using lower-fidelity simulators.
- Furthermore, we are uncertain about the generalizability to other levels of automation and different contexts using the same automated driving system.

Example II: Traffic sign recognition

- An **evaluation study** was performed at DLR in Brunswick, utilizing the MoSAIC-Lab **driving simulator**, comparing three different display options and their influence on **behavioural validity**



- In order to analyse **absolute validity**, a small **real world pre-study** was conducted, analyzing the traffic sign readability capabilities in reality and in simulation

Example II: Pre-study results

Study set-up

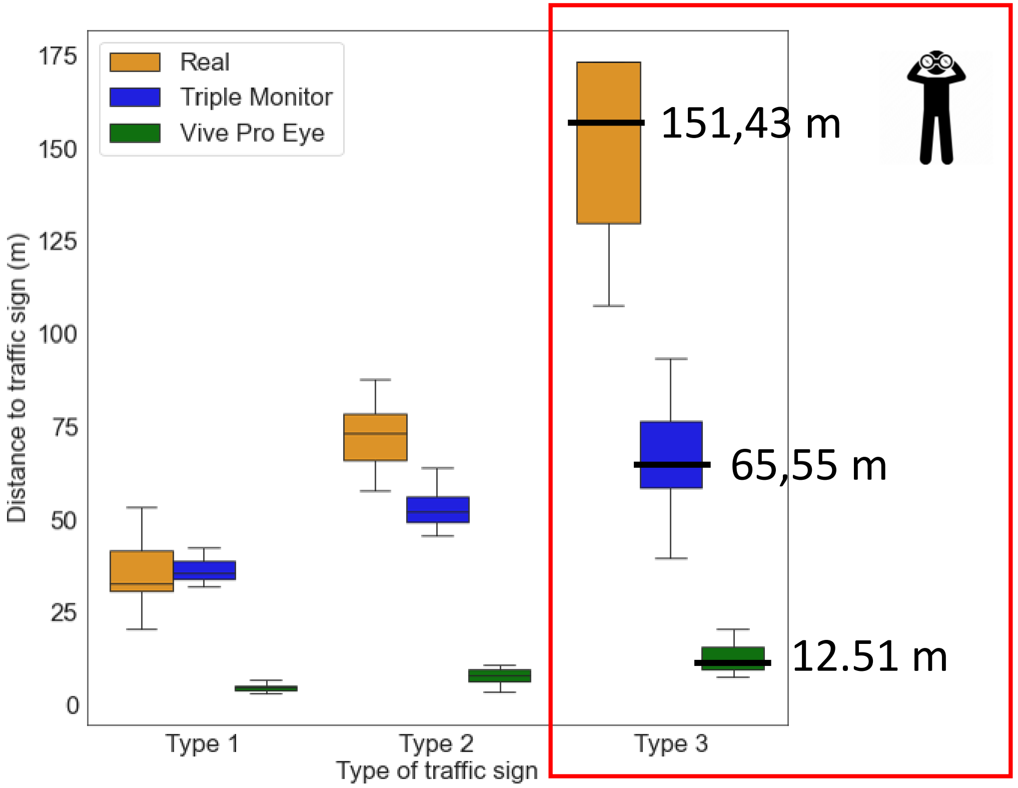
- Reality
  - Approaching traffic sign
    → Report, when traffic sign is readable


- Simulation
  - Approaching traffic sign
    → Press button when traffic sign is readable
  - Static position, fixed distance to traffic sign
    → Scale size of traffic sign until readable
  - 3 types of traffic signs:

1 2 3

Absolute Validity

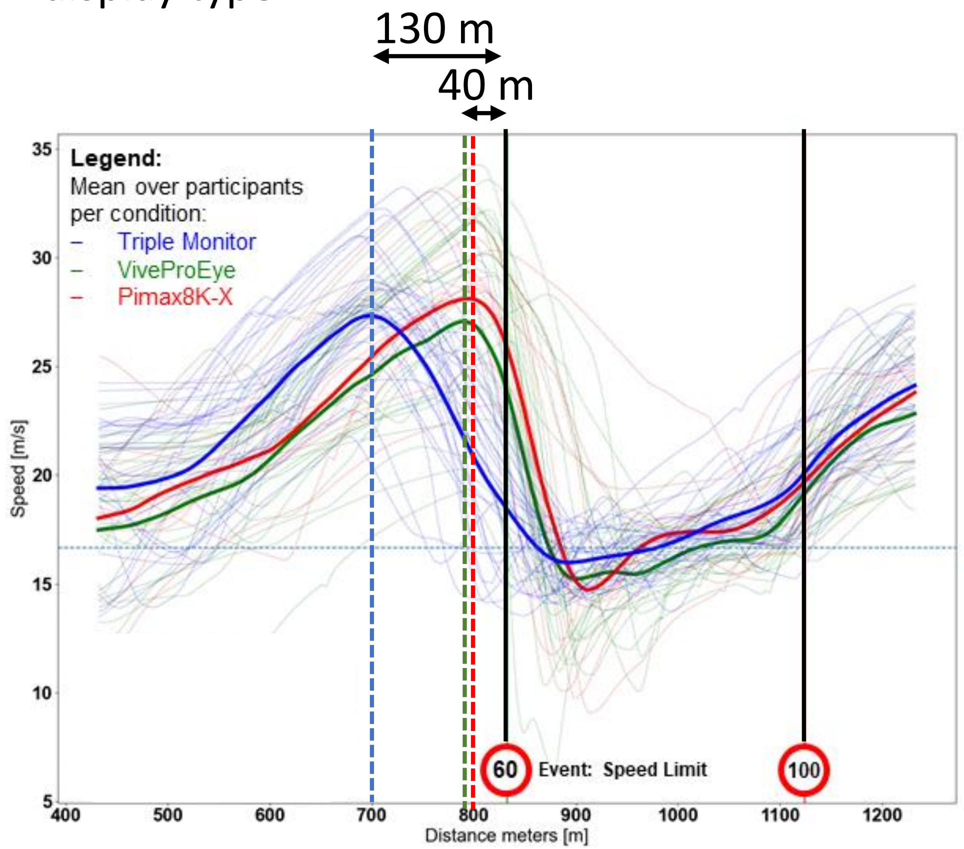
- Readability is strongly dependent on display type and sign type



Example II: Main study results – behavioral validity

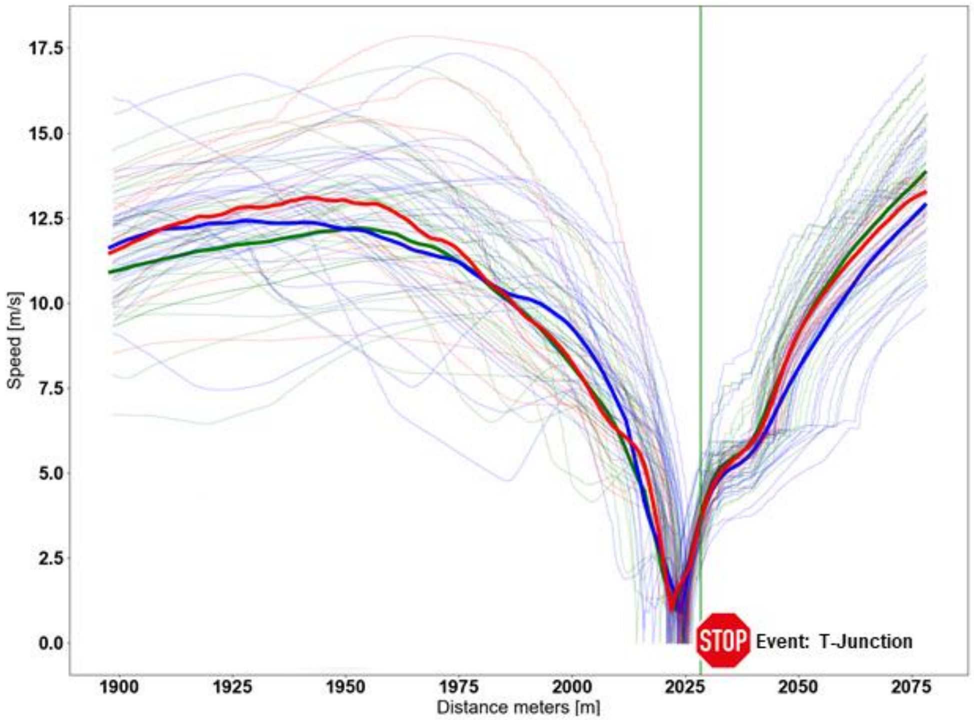
Scenario: Approaching Speed Limit

- Significant behavioral difference dependent on display type



Scenario: Approaching Stop Sign

- No Significant behavioral difference at stop sign
- Traffic sign readability is not the main factor for braking decision



General discussion

- When planning validation studies, consider your use case/typical simulator usage, as driving simulator validity depends on the **use case**.
- Model **verification** should be undergone prior to testing in the simulator.
- Adequate **statistical methodology** should be used, which is capable of supporting the underlying hypotheses.
- The **credibility** of a simulator study's results should be derived based on the available validation and verification results, taking into account differences in the study settings.
- **Uncertainty** may result from a lack of transferable validation studies, or the incapability of performing validation. The degree of acceptable uncertainty depends on the **risk** associated with decisions based on simulator studies.

DSC 2024 EUROPE VR

Driving Simulation & Virtual Reality Conference & Exhibition

18-20 September 2024

Palais des Congrès et de la Musique, Strasbourg | France

THANK YOU