Assessing the validity of a driving simulation for comparing in-vehicle informational interfaces

Ying Wang, Kathryn Godfrey, Bruce Mehler, Bryan Reimer & Joseph Coughlin

University Transportation Center

AGELAB Life Tomorrow

Abstract

Validity is very important in deciding the applicability of driving simulation, as a tool providing safe, controlled, and replicable research protocols, for various research and design studies. Data from on-road and simulation studies were compared to assess the validity of measures generated in the simulator. In the on-road study, driver interaction with three manual address entry methods (touch screen, key pad, and scroll wheel) was assessed in an instrumented vehicle to evaluate relative usability and safety implications. A separate group of participants drove a similar protocol in a medium fidelity, fixed-base simulator to assess the extent to which simulator measures mirrored those obtained in the field. Visual attention and task measures mapped very closely between the two environments. In general, however, driving performance measures did not differentiate among devices at the level of cognitive demand employed in this study. The findings obtained for visual attention and task engagement suggest that medium fidelity simulation provides an effective means to evaluate the effects of IVIS designs on these categories of driver behaviour.

Background

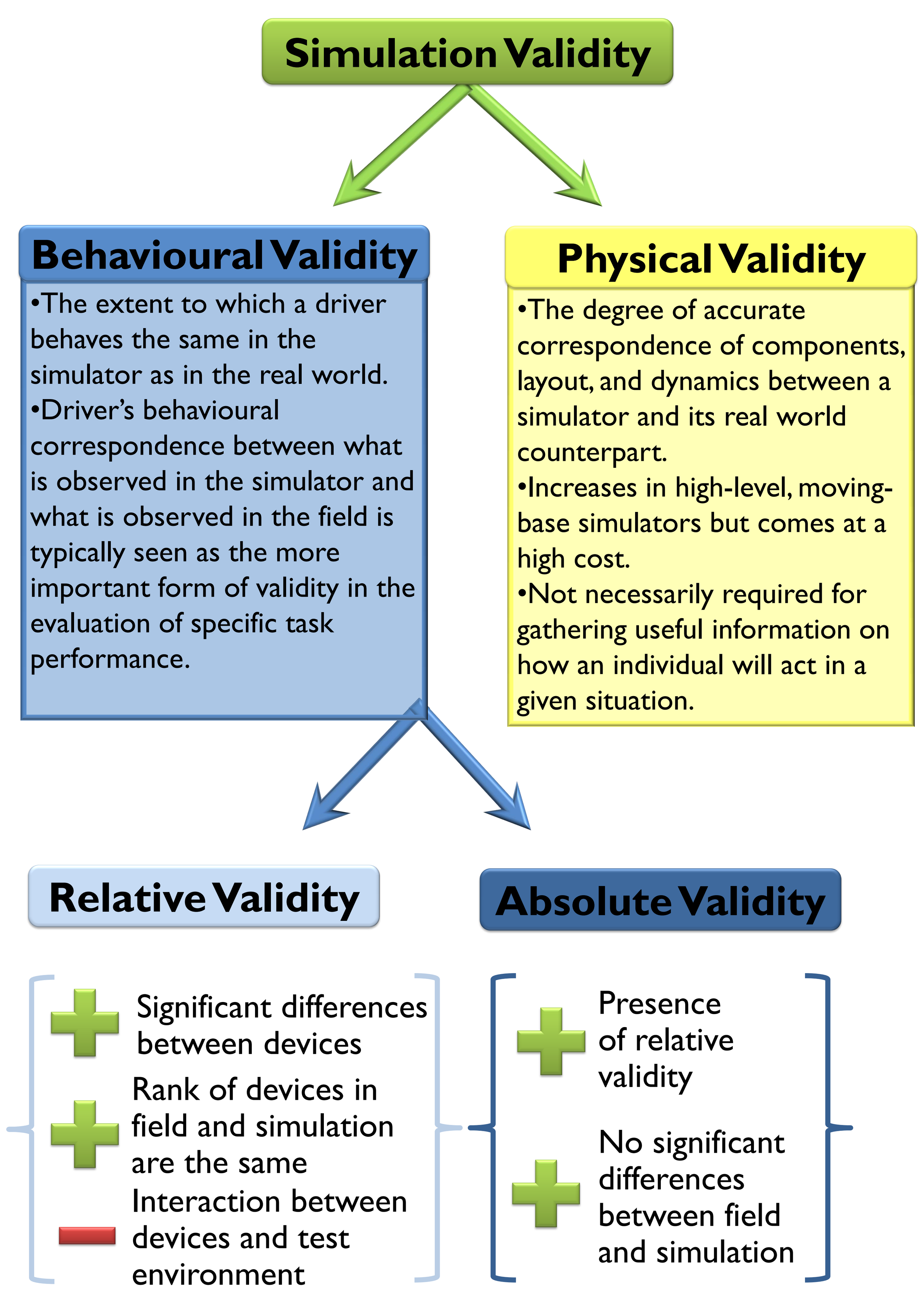


Figure 1. Theoretical framework of simulation validity and the conditions in this experiment that either support (green plus sign) or oppose (red minus sign) relative and absolute validity.

Methods

Participants

- Age: 22-28
- Driving experience: > 3yr
- On-road: 28 participants
- Simulation: 30 participants

Design

- Between group factor
  - Environment (env, 2 Level) On-Road, Simulator
- Within group factor
  - Input device (dev, 3 Level) Touch Screen Key Pad Scroll Wheel

Apparatus

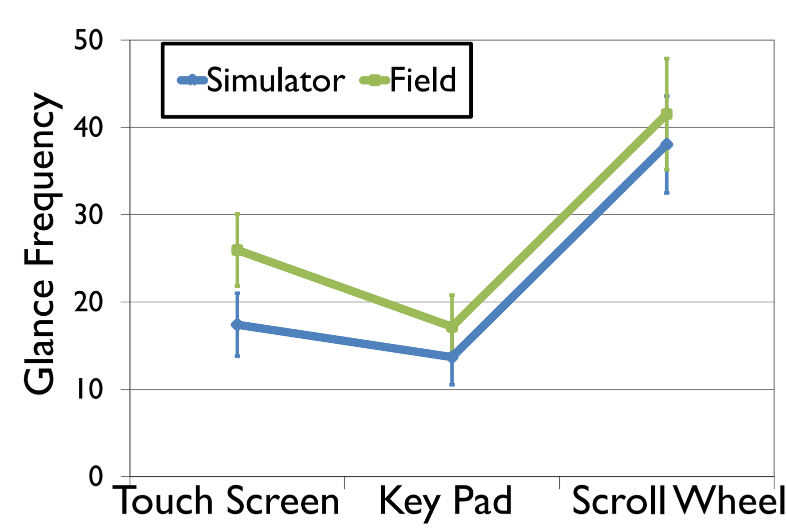
- On-Road: MIT AgeLab "Aware Car"
- Simulation: "Miss Daisy" Simulator
- Eye tracking: Seeing Machines FaceLab 4.2
- Surrogate in-vehicle information systems

Destination Entry Task

- Each include a state, a city and a street
- Entry required only the 1st & 2nd letter of each word, remaining letters were auto-completed
- 3 consecutive repetitions for each device
- Randomized input device presentation order

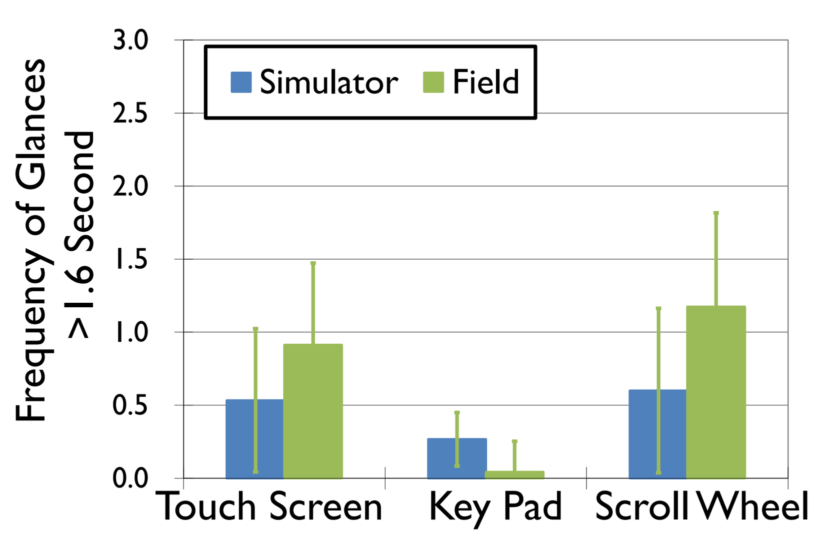
Results

Visual Attention

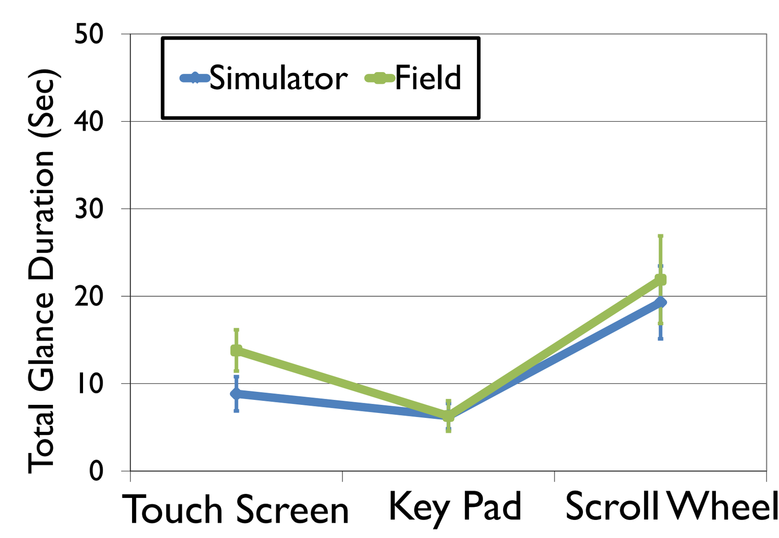


Graph 1. Glance frequency during destination entry task (P_{dev} <.001, P_{env} <.05, $P_{(dev \times env)}$=.231). Error bars represent 95% confidence interval.

Figure 2. Eye Tracking (top), Scroll Wheel (left), Touch Screen and Key Pad (right)

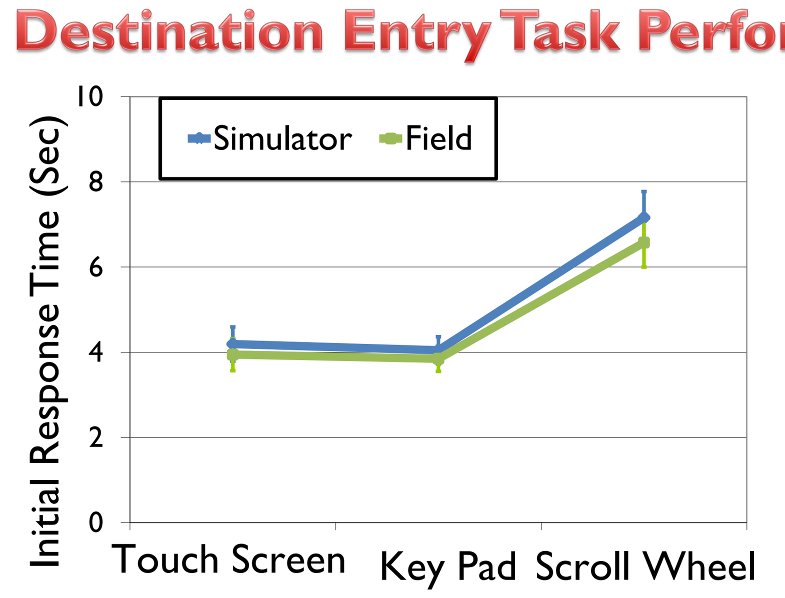


Graph 2. Frequency of glances >1.6 seconds during destination entry task (P_{dev} <.01, P_{env}=.299, $P_{(dev \times env)}$=.147. Error bars represent 95% confidence interval.

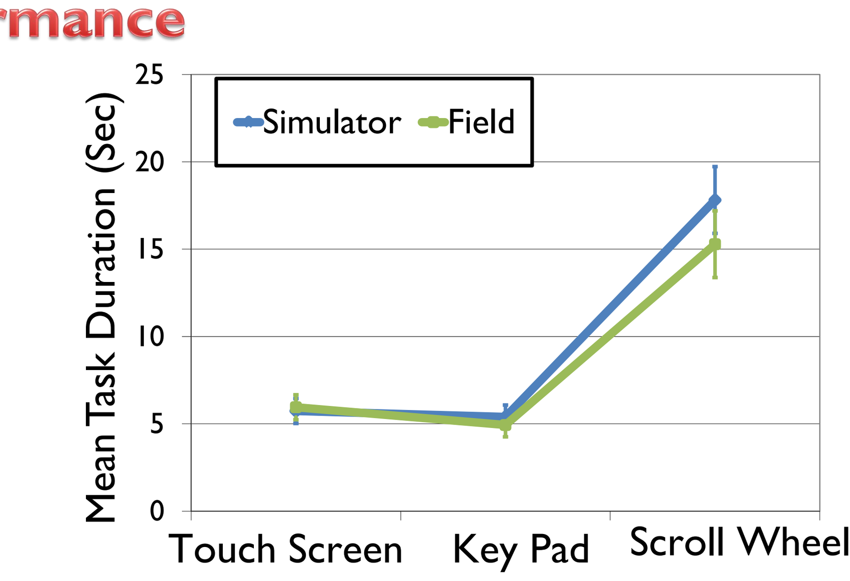


Graph 3. Total glance duration during destination entry task (P_{dev} <.001, P_{env}=.126, $P_{(dev \times env)}$=.142. Error bars represent 95% confidence interval.

Destination Entry Task Performance

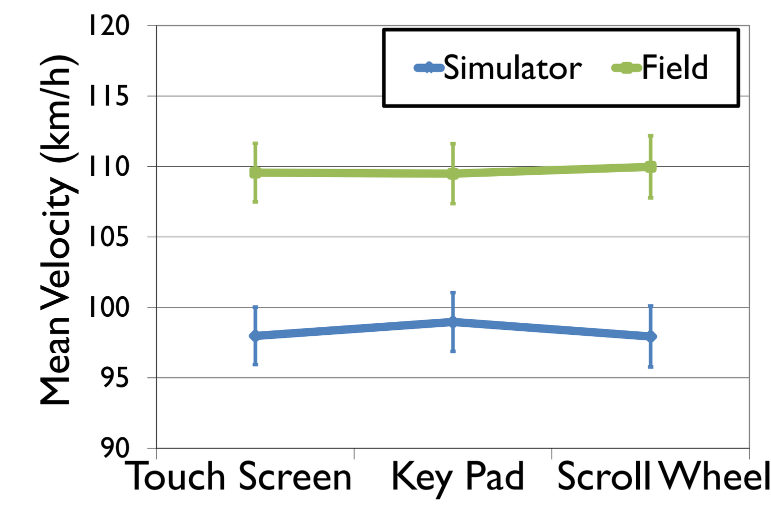


Graph 4. Initial response time during destination entry task (P_{dev} <.001, P_{env}=.175 $P_{(dev \times env)}$=.453. Error bars represent 95% confidence interval.



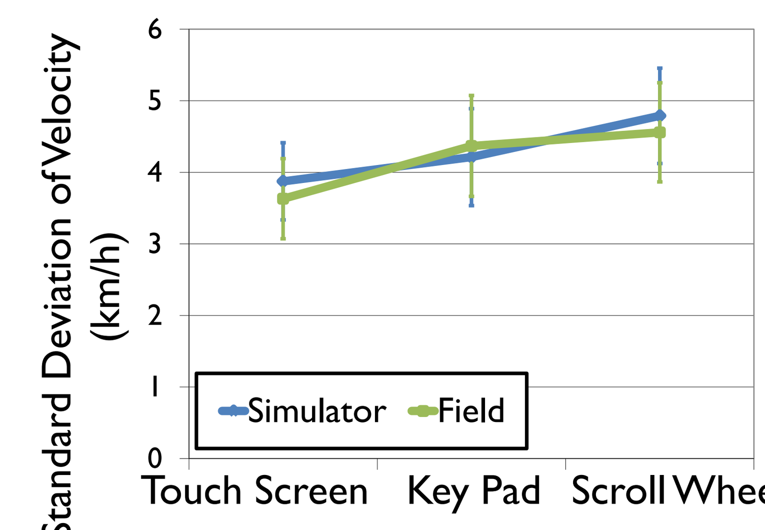
Graph 5. Mean duration of destination entry task (P_{dev} <.001, P_{env}=.108, $P_{(dev \times env)}$=.070. Error bars represent 95% confidence interval.

Driving Performance

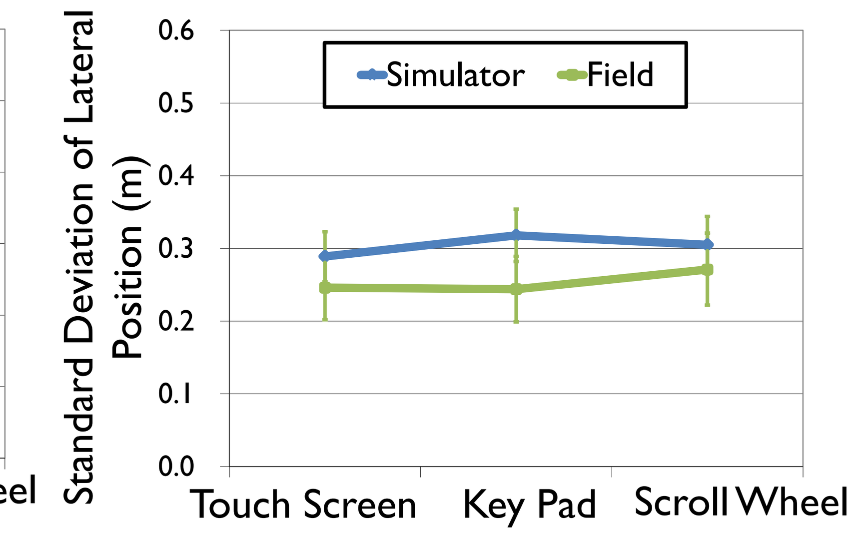


Graph 6 Mean velocity during destination entry task (P_{dev}=.748, P_{env}<.001, $P_{(dev \times env)}$=.432. Error bars represent 95% confidence interval.

Figure 3 MIT AgeLab Aware Car and Miss Daisy Simulator



Graph 7 Standard deviation of velocity during destination entry task (P_{dev} <.01, P_{env}=.761, $P_{(dev \times env)}$=.668. Error bars represent 95% confidence interval.



Graph 8 Standard deviation of lane position during destination entry task (P_{dev}=.416, P_{env}<.05, $P_{(dev \times env)}$=.424. Error bars represent 95% confidence interval.

Establishment of Validity

	Measurements	Rel. validity	Abs. validity
Visual Attention	Glance frequency	+	−
	Total glance duration	+	+
	Frequency of glances >1.6s	+	+
Destination Entry Task	Initial response time	+	+
	Mean task duration	+	+
Driving	Mean forward velocity	*	−
	Std. dev. forward velocity	+	+
	Std. dev. lane position	*	−

Conclusion

- Visual attention and secondary task performance measures appear particularly promising for modeling the effects of on-road drivers' interactions with an in-vehicle information systems interface.
- Measures of glance frequency, total glance duration, initial response time, and mean task time mapped almost identically between simulation to field.
- Compared to standard driving performance measures (mean velocity and standard deviation of lane position), the visual attention measures appear more sensitive for detecting subtle differences between HMI designs.
- Standard deviation of forward velocity was the only driving performance measure to meet criteria for both relative and absolute validity in this study, and the statistical significance was modest.
- In conclusion, fixed-based driving simulation appears an acceptable method of modeling basic task performance and visual distraction, but not driving performance measures.

Acknowledgements

Funding for this project was provided in part by the United States Department of Transportation's Region I New England University Transportation Center at the Massachusetts Institute of Technology, the Santos Family Foundation, and indirect support was provided by Ford Motor Company through the use of the instrumented vehicle. The authors wish to acknowledge the contributions of Jonathon Long and Alexander Pina for their contribution in the collection of data and preparation of this manuscript.