13

Simulator Validity: Behaviors Observed on the Simulator and on the Road

Nadia Mullen
Lakehead University

Judith Charlton
Monash University

Anna Devlin
Monash University

Michel Bédard
Lakehead University

Abstract

The Problem. Driving simulators offer a safe, convenient alternative to measuring driving performance on-road. However, the results of simulator studies may not generalize to driving in the real world if the simulator lacks behavioral validity. Behavioral validity refers to the extent to which the simulator elicits the same driving behaviors that occur when driving in the real world. ***Role of Driving Simulators.*** Validation is important to generate and maintain simulator use, acceptance, and credibility, and is vital when simulator performance influences real-world outcomes, such as road or vehicle design, or whether drivers retain their license. ***Key Results of Driving Simulator Studies.*** A review of studies evaluating the behavioral validity of simulators showed that simulators provide a valid tool for assessing a variety of driving performance measures such as speed, lateral position, brake onset, divided attention, and risky traffic behaviors. Simulators also appear sensitive to age-related changes in driving performance and cognition. Measures for which simulators do not appear valid are discussed, in addition to factors influencing validity, such as driving ability. Overall, the evidence reviewed in this chapter indicates that simulator driving behavior approximates (relative validity), but does not exactly replicate (absolute validity), on-road driving behavior. This is sufficient for the majority of research, training, and assessment purposes for which simulators are used. However, where absolute values are required, on-road measures will generally be necessary. ***Scenarios and Dependent Variables.*** Validation studies involve consideration of factors such as the research question, task conditions, and dependent measures, each of which can affect validity. We discuss these methodological considerations, as well as statistical techniques used to establish validity. ***Platform Specificity and Equipment Limitations.*** Assumptions about driving simulator validity are critically dependent on the specific experimental conditions under which the driving behaviors are compared. Variations across simulator equipment, software, and environment may affect the generalizability of validation results. Therefore, each simulator should be validated for its ability to measure the driving behavior of the cohort for which it is to be used.

13.1 Introduction

This chapter addresses the topic of the behavioral validity of simulators, that is, the extent to which the simulator induces the same driving behaviors that occur when driving in the real world. The issue of validation is important in terms of generating and maintaining simulator use, acceptance, and credibility. It is especially vital when simulator performance outcomes are used to influence real-world outcomes, such as road or vehicle design, or whether drivers retain their license.

Driving simulator validation studies have generally assessed validity in terms of physical and behavioral validity (Blaauw, 1982; Blana, 1996). *Physical validity* refers to the extent to which the physical components of the simulator vehicle correspond to the on-road vehicle, including the simulator layout, visual displays, and dynamics such as the feel of braking and steering controls (see Table 13.1, and see in this book, chap. 8 by Andersen and chap. 7 by Greenberg & Blommer). *Behavioral validity* refers to the level of correspondence between the driving behaviors elicited in the simulator and on real roads (see also Ranney, this book, chap. 9). In some contexts, both physical and behavioral validity have been used to describe the simulator's *external validity* (Blana & Golias, 2002; Klee, Bauer, Radwan, & Al-Deek, 1999; Reed & Green, 1999; Törnros, 1998) or the extent to which driving behavior in the simulator can be generalized to real driving behavior (see Table 1; Hoskins & El-Gindy, 2006).

Behavioral validity has been further defined in terms of *absolute* and *relative validity* (Blaauw, 1982). While absolute validity requires that the two driving environments produce the same numerical values, relative validity is established when the differences between the two environments are in the same direction, and of the same or similar magnitude (see Table 1; Godley, Triggs, & Fildes, 2002). Absolute validity is rarely established in driving simulator studies due to a number of driving simulator characteristics that are not found in real on-road settings. Indeed, it is argued that absolute validity is not necessary for simulators to be useful research tools; rather, establishing relative validity is necessary and sufficient (Törnros, 1998). This is particularly so when the researcher is interested in comparing changes in driving patterns under different treatment conditions across simulator and real-world settings. In these circumstances, it is more important to establish that the treatment resulted in the same kind of behavior change (e.g., speed reduction) compared with a control condition, in both settings, and less critical to establish that the magnitude of change (e.g., amount of speed reduction) is precisely matched across settings.

13.2 Methodological Considerations

13.2.1 The Research Question, Task Conditions, and Dependent Measures

The concept of simulator validity has relevance only in the context of a specific research question and it is likely that the level of similarity of driving behaviors across simulator and real driving settings may be highly task-dependent. Participant characteristics may also influence validity, such as participants' experience of simulator discomfort and driver motivation (Blana, 1996), driver age and experience, and the presence of medical conditions. That is to say, while a simulator may be described as a valid instrument for training novice drivers to undertake hazard-avoidance maneuvers in complex city traffic, the same device may not be valid for evaluating the performance of experienced drivers or drivers with impairments on a different task, such as highway driving with little or no traffic and no hazardous events. Indeed, Kaptein and colleagues caution that "any use of driving simulators should be preceded by questioning whether the simulator is sufficiently valid for the task or ability to be investigated" (Kaptein, Theeuwes, & van der Horst, 1996, p. 31).

Assumptions about driving simulator validity are also critically dependent on the specific experimental conditions under which the driving behaviors are compared. Variations across simulator equipment (e.g., number and size of display monitors), software (e.g., simulator program), and environment (e.g., temperature, noisiness) may affect the generalizability of validation results. Several simulator-based factors have been implicated in influencing simulator validity, including the fidelity of proprioceptive information, the presence of a motion platform, and the quality of image resolution of the display of road and traffic environment (Kaptein et al., 1996).

Similarly, the operational definition of real-world driving (i.e., the precise conditions under which real-world driving performance is measured) can influence the authenticity of validation data. Validation studies (i.e., studies that compare real-world with simulated performance) reviewed in this chapter differed widely in their operational definitions of real-world driving. Some studies examined on-road driving in naturalistic or "uncontrolled" conditions (e.g., driving in real traffic with a standard vehicle); in other studies, on-road driving was examined in more controlled conditions (e.g., driving an instrumented vehicle on a test track with an experimenter as a passenger; Blana, 1996).

Assessments of simulator validity may produce widely varying results, depending on which driving measures are compared. For example, high levels of similarity for one measure of driver behavior indicating validity across simulator and real-world settings, such as for speed, may not necessarily be observed for other measures, such as braking, steering, or lane position. Kaptein et al. (1996) emphasized the importance of selecting measures to adequately assess simulator validity; not all variables measured during a simulator drive accurately reflect the same measure in an on-road drive. In a study involving older drivers, Lee, Cameron and Lee (2003) correlated age with individual behavior measures obtained from the simulator to establish criterion-related validity (i.e., to check that each simulator measure showed expected age-related declines in driving performance) prior to analyzing the variables of interest. A key point here in relation to study design is that not all measures may be necessary or relevant for the task and research question under investigation.

Variations in data collection methods may play a role in influencing data accuracy and authenticity (Blana, 1996). For example,

TABLE 13.1 Examples of Evaluating Validity

Type of Validity	Source of Definition	Possible Methods of Testing for this Validity	Examples and Expected Results when Validity is Established[a]
Behavioral Validity			
Absolute validity (established when the simulated and on-road drives produce the same numerical values)	Blaauw, 1982	ANOVA	*Example:* Speed measured during a simulated and an on-road drive is compared. *Expect:* Means will not differ significantly. *e.g.,* mean simulator speed = 50 km/h, mean on-road speed = 53 km/h, $p > 0.05$
		Visual inspection + correlation	*Expect:* Graphed data will have similar shape and may overlap, and correlation will be significant.
Relative validity (established when the simulated and on-road drives produce numerical values similar in magnitude and in the same direction)	Törnros, 1998	ANOVA	*Example:* Speed measured during a simulated and an on-road drive, on both urban and rural roads, is compared. *Expect:* Means may differ significantly between the simulated and on-road drive, but they will be in the same direction on each level of the independent variable, and of similar magnitude. *e.g.,* Urban roads: Mean simulator speed = 50 km/h, mean on-road speed = 58 km/h, $p < 0.05$ Rural roads: Mean simulator speed = 80 km/h, mean on-road speed = 89 km/h, $p < 0.05$
		Visual inspection + correlation	*Expect:* Graphed data will have similar shape and be non-overlapping, a similar distance will separate the lines at each level of the independent variable, and correlation will be significant.
Interactive absolute validity (established when the simulated and on-road drives produce the same numerical values over time)	Godley et al., 2002	ANOVA	*Example:* Speed measured at 5 points along the same 50 m section of a simulated and an on-road drive is compared. *Expect:* Means at each of the 5 points will not differ significantly. *e.g.,* Point 1: Mean simulator speed = 50 km/h, mean on-road speed = 53 km/h, $p > 0.05$ Point 2: Mean simulator speed = 55 km/h, mean on-road speed = 59 km/h, $p > 0.05$ Point 3: Etc.
		Visual inspection + correlation	*Expect:* Graphed data will have similar shape and may overlap at all time points, and correlation will be significant.
Interactive relative validity (established when the simulated and on-road drives produce numerical values similar in magnitude and in the same direction over time)	Godley et al., 2002	ANOVA	*Example:* Speed measured at 5 points along the same 50 m section of a simulated and an on-road drive is compared. *Expect:* Means at each of the 5 points may differ significantly between the simulated and on-road drive, but they will be in the same direction at each time point, and of similar magnitude. *e.g.,* Point 1: Mean simulator speed = 50 km/h, mean on-road speed = 58 km/h, $p < 0.05$ Point 2: Mean simulator speed = 55 km/h, mean on-road speed = 64 km/h, $p < 0.05$ Point 3: Etc.
		Visual inspection + correlation	*Expect:* Graphed data will have similar shape and be non-overlapping at all time points, a similar distance will separate the lines at each time point, and correlation will be significant.

(continued)

TABLE 13.1 (Continued) Examples of Evaluating Validity

Type of Validity	Source of Definition	Possible Methods of Testing for this Validity	Examples and Expected Results when Validity is Established[a]
Physical Validity			
Physical validity (the extent to which the physical components of the simulator correspond to on-road vehicles)	Blaauw, 1982	Visual inspection of the simulator components and layout, and its dynamic characteristics such as the response of braking and steering controls	*Example:* The physical features of a simulator are compared with a real vehicle. *Expect:* A high degree of similarity between the simulator and on-road vehicle. *e.g.*, A simulator with a full cab, motion-base, and projected 360 degree view has higher physical validity than a simulator with a car seat only, fixed-base, and view presented on a 17" computer monitor.
External Validity			
External validity (the extent to which the results obtained in the simulator can be generalized to driving on the road; can include physical and behavioral validity)	Kaptein et al., 1996	ANOVA; Visual inspection + correlation	Combination of behavioral and physical validity examples.
Internal (Criterion) Validity			
Internal validity (the extent to which causal inferences about the impact of an experimental treatment [e.g., the introduction of speed limits resulting in a speed reduction] can be made with confidence; can include physical and behavioral validity)	Reimer et al., 2006	ANOVA; Visual inspection + correlation	Combination of behavioral and physical validity examples.

[a] Significant differences mentioned in this column refer to statistical significance and do not imply practical significance.

some simulator measures are generated by the computing facility of the simulator, while other measures rely on observational techniques. In some research contexts, it may be pertinent to establish consistency across these different methods of data collection. For example, using a STISIM 400 simulator, Bédard and colleagues established a significant correlation ($r = 0.83$) between simulator-recorded driving errors and real-time researcher-recorded demerit points in the same simulator setting (Bédard, Parkkari, Weaver, Riendeau, & Dahlquist, 2010). They also showed a high level of consistency across driver behavior measures conducted in real-time with scores determined from later video analyses ($r = 0.83$). These results suggest that for the driving errors considered in Bédard et al. (2010) study, the various methods of collecting data did not affect the simulator's validity. However, as noted above, these findings may not generalize to other measures in this simulator (e.g., speed, braking responses) or to measures of driving errors in other simulators because any difference (e.g., in equipment, surroundings, or measurement technique) could affect validity. Furthermore, the relationships between the measures reported here for simulator driving behaviors will not necessarily hold for real-world driving; direct comparison of measures across the two settings would be required to determine this.

13.2.2 Analysis

Typically, simulator validation studies have used conventional statistical approaches for group comparisons, such as analysis of variance (ANOVA), to examine differences between simulator and on-road driving performance measures (Klee et al., 1999; McGehee, Mazzae, & Baldwin, 2000; Reed & Green, 1999). For example, absolute validity is claimed if group means for a driving behavior (e.g., speed) measured during the simulated and on-road drives do not significantly differ (Table 13.1). If mean performance measures differ significantly between the simulated and on-road drives, but are in the same direction across levels of another independent variable (e.g., speeds on urban roads being slower than speeds on rural roads during both simulated and on-road drives), relative validity will have been established (Table 13.1). These statistical procedures have been used to establish both absolute and relative validity for key driving variables such as mean speed, braking response time, variability in lane position, or steering wheel angle, measured at specific points of interest in the drive or expressed as an average across a designated section of the drive. Table 13.2 provides a summary of the studies reviewed in the following section of this chapter, and presents information concerning the statistical approach taken to assess simulator validity for each driving measure.

Another approach for evaluating validity relies on a comparison of specific aspects of the driving profile (such as speed or error measures) for a key region of a drive or across an entire drive, determining a simulator's *interactive* (or *dynamic*) *relative validity* (see Table 13.1). This approach captures the collective dynamic nature of driving and allows the comparison of patterns of performance *throughout* a driving task, in drivers responding

TABLE 13.2 Summary of Driving Simulator Validation Studies Cited in the Chapter

Validation Studies	Driving Simulator	Data Collection Technique	n/Age	Statistical Approach	Comments/Key Findings: Absolute Validity Established	Comments/Key Findings: Relative Validity Established
Bédard, 2008	STISIM	Sim. and on-road; physiological responses	$N = 24$ 21–57 yrs	ANOVA		Perceived threat from unexpected events
Bédard et al., 2010	STISIM	Sim. and on-road; cognitive performance	$N = 8$-38 67–81 yrs	Correlation		Sim. driving and on-road demerit points
Bella, 2005	Inter-Uni. Research Center for Road Safety (CRISS)	Sim. and on-road	$N = 35$ 24–45 yrs	Comparison of means—bilateral Z test for non-matched samples		Speed
Bella, 2008	CRISS	Sim. and on-road	$N = 40$ 23–60 yrs	Comparison of means—bilateral Z test for non-matched samples	Speed for non-demanding road configurations	Speed and complexity of maneuver
Bittner et al., 2002	Uni. of Washington Real Drive	Sim. and on-road	$N = 12$ 21–34 yrs	Generalized Youden plots		Curve entry speeds
Blaauw, 1982	Institute for Perception TNO	Sim. and on-road	$N = 48$ 18–36 yrs	ANOVA, correlation	Speed	Lateral displacement
Blana & Golias, 2002	Leeds Advanced Driving Simulator (LADS)	Sim. and on-road	$N = 100$ M age = 36 yrs	Independent samples t test		No – lateral displacement
Charlton et al., 2008	Monash Uni. Accident Research Centre (MUARC)	Sim. and on-road	$N = 30$ 29–89 yrs Vision impaired and controls	ANOVA, correlation		Speed and gaze direction
Fisher et al., 2007	Uni. of Massachusetts-Amherst advanced fixed-base driving simulator	Sim. and on-road	$N = 12$-24 18–21 yrs	Independent samples t test		Training effects and gaze direction
Godley et al., 2002	MUARC	Sim. and on-road, speed countermeasures	$N = 24$ 22–52 yrs	ANOVA; correlations and canonical correlations		Speed countermeasures
Hakamies-Blomqvist et al., 2001	VTI	Sim. and on-road	$N = 35$ 66–80 yrs	ANOVA		Steering wheel angle and lane position
Hoffman et al., 2002	Iowa Driving Simulator (IDS)	Sim. and on-road	$N = 16$ 25–55 yrs	Descriptive statistics: Comparison of means		Mean braking onset
Klee et al., 1999	Uni. of Central Florida (UCF)	Sim. and on-road	$N = 30$ 17–65 yrs	A bilateral Z test for non-matched samples		Forward speed
Lee, Cameron, et al., 2003	STISIM	Sim. and on-road	$N = 129$ 60–88 yrs	Correlations and principal component analysis		Sim. driving performance and on-road driving performance
Lee et al., 2007	STISIM	Sim. and on-road	$n = 50$ PD patients $n = 150$ controls 60–80 yrs	Pearson correlation, stepwise linear regression		Control participants' overall driving performance in sim. and on-road
Lee et al., 2005	STISIM	Sim., cognitive performance, driver violation points 2 year interval	$N = 129$ 65–88 yrs	Hierarchical Poisson regression analysis		The prediction of older driver crash history from driving sim. performance
Lee, Lee, & Cameron, 2003	STISIM	Sim. and measure of visual attention	$N = 129$ 60+ yrs	Repeated measures ANOVA		Visual attention skill and age

(continued)

TABLE 13.2 (Continued) Summary of Driving Simulator Validation Studies Cited in the Chapter

Validation Studies	Driving Simulator	Data Collection Technique	n/Age	Statistical Approach	Comments/Key Findings: Absolute Validity Established	Comments/Key Findings: Relative Validity Established
Lee, Lee, Cameron, & Li-Tsang, 2003	STISIM	Sim. and crash history	$N = 129$ 60–88 yrs	Logistic regression		Driving performance and previous crash history
Lew et al., 2005	STISIM	Sim. and on-road, observational data	$n = 11$ TBI $n = 16$ controls 18–58 yrs	Correlations		No - TBI individuals' driving performance
McAvoy et al., 2007	Doron Precision Systems Inc. AMOSII	Sim. and on-road	$N = 127$ 18-70 yrs	ANOVA		No - speed through a work zone
McGehee et al., 2000	IDS	Sim. and on-road	$N = 120$ 25-55 yrs	ANOVA and 95% confidence interval	Total brake reaction time	
Philip et al., 2005	Divided Attention Steering Simulator	Sim. and on-road; self-rated fatigue; reaction time	$N = 12$ 19–24 yrs	ANOVA		Inappropriate line crossings due to fatigue
Reed & Green, 1999	Uni. of Michigan Transportation Research Institute (UMTRI)	Sim. and on-road and a concurrent manual dialling phone task	$N = 12$ 20–30 yrs and 60+ yrs	ANOVA	Speed control	Effects of a phone task and age
Reimer et al., 2006	Instrumented full cab 2001 Volkswagen Beetle	Validated sim. with respect to self-report questionnaires	$N = 48$ 16–55 yrs	A multi-trait-multi method correlation matrix of driving behavior		Self-reported survey items and sim. performance
Riemersma et al., 1990	Daimler-Benz	Sim. and on-road	$N = 24$ 20–35 yrs	Descriptive statistics: Comparison of a reduction in mean speed		Speed reduction measures
Shinar & Ronen, 2007	STISIM	Sim. and on-road	$N = 16$ 28–30 yrs	Correlation and regression analysis		Changes in speed
Slick et al., 2006	DriveSafety DS-600c motion-based simulator	Sim. and on-road; physiological responses	$N = 22$ M age = 16 yrs	ANOVA	Physiological responses during turns at intersections	Physiological responses during entire drive
Törnros, 1998	VTI driving sim.	Sim. and on-road through a tunnel	$N = 20$ 23-52 yrs	ANOVA		Speed and lateral position
Toxopeus, 2007	STISIM	Sim.	$N = 79$ 18–30 yrs and 60–80 yrs	ANOVA		Reaction time (to brake and press horn) and age; speed and age
Volkerts et al., 1992	TS2 driving simulator	Sim. and on-road	$N = 18$ 25–31 yrs	MANOVA, correlation and multiple regression		No – sedative drug induced impairment
Wade & Hammond, 1998	Wrap around sim. (WAS) Uni. of Minnesota	Sim. and on-road	$N = 26$ 18–34 yrs	Descriptive statistics: Raw data charts of lane position from centerline		Lane position
Watts & Quimby, 1979	Transport and Road Research Laboratory	Sim. and on-road; physiological responses and perceived hazard risk ratings	$N = 60$ Age range unknown	Correlation coefficient	Perceived risk of hazards on the road and in the sim.	
Yan et al., 2008	UCF	Sim. and on-road; crash history	$N = 241$ 15–45+ yrs	ANOVA		Risky behaviors at signalized intersections

to different events or treatments. Godley et al. (2002) investigated interactive validity by comparing patterns of driver speed control in a simulator and on a real road, in response to rumble strips. In addition to their examination of absolute and relative validity, using conventional measures, the authors examined speed profiles over the entire data collection area (derived from one meter averages) and applied a correlation analysis, based on canonical correlation, to evaluate patterns of speed management in response to treatment and control (no rumble strip) sites on the road and in the simulator (shown in Figure 13.1, respectively). Participants drove faster at both treatment and control sites in the simulator compared with the on-road setting (arguing against absolute validity). However, when averaged over the total data collection area and across both experiments, speed at the treatment site stop sign approach was significantly slower than at the control site. Treatment site speeds were slower in both the simulator and on-road settings. Thus, relative validity was established for the stop sign approach speed. Moreover, the authors noted that the *pattern of speeds for the treatment site relative to the control site* was also similar in both simulator and on-road settings throughout most of the stop sign approach, as demonstrated by a significant correlation ($r = 0.40$), suggesting support for interactive relative validity.

Finally, as is evident from Table 13.2, the association between the on-road and simulator data when establishing interactive relative validity need not be confined to a linear one. Validity is established when any mathematical function significantly relates the on-road and simulator data. This type of validity, which may be thought of as *functional validity*, allows for consideration of a range of mathematical relationships between the variables, including both linear and non-linear associations. Functional validity is a type of behavioral validity, and linear associations are synonymous with absolute or relative validity. However, there are many types of functions that may explain

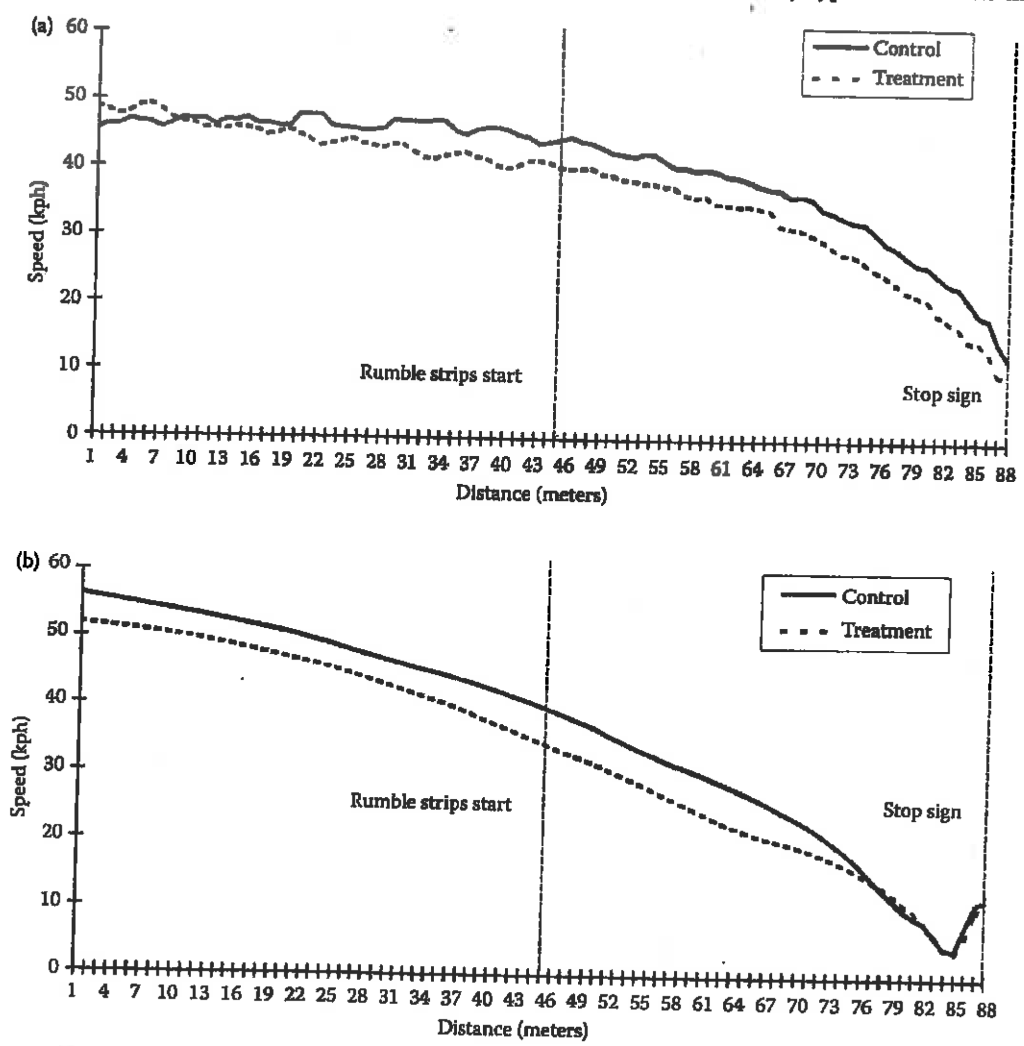

FIGURE 13.1 Mean approach speed to a stop sign in the (a) on-road drive, and (b) simulator drive. (Godley, S. T., Triggs, T. J., & Fildes, B. N. (2002). Driving simulator validation for speed research. *Accident Analysis & Prevention*, *34*, 589–600. Copyright 2002 by Elsevier Science. Reprinted with permission.)

the relationship between on-road and simulator data (e.g., linear, quadratic, exponential, hyperbolic), not all of which may indicate an adequate level of validity.

13.3 Review of Studies Investigating Behavioral Validity

Despite the importance that is placed on the simulator as a tool for understanding driving performance and driver competence, an extensive literature search revealed a limited number of studies specifically evaluating the behavioral validity of simulators (Table 13.2). Most of the available literature on simulator validation has focused on measures of speed, lateral position, and braking responses, with a few studies examining more complex driving behaviors (e.g., performance while dialing a car phone) or behaviors of specific driver groups (e.g., older drivers). The following section explores the literature on this topic with a particular emphasis on the selection of performance measures, the tasks/conditions under investigation, and the evidence provided for behavioral validity of specific simulators and simulated driving behaviors.

13.3.1 Speed

One of the most commonly studied measures of behavioral validity of driving simulators is driver speed (Bella, 2005; Bella, 2008; Bittner, Simsek, Levison, & Campbell, 2002; Blaauw, 1982; Blana, & Golias, 2002; Godley et al., 2002; Klee et al., 1999; Törnros, 1998). Overall, studies have consistently demonstrated relative (but not absolute) validity for speed data. Research by Godley et al. (2002), presented above, shows strong evidence for both relative and interactive relative validity for speed on the Monash University Accident Research Centre (MUARC) driving simulator, but absolute values of speed differed across the simulated and real-world driving settings. Similarly, recent work by two of the authors using the same simulator has shown preliminary evidence for relative coupling of speed profiles in a simulator and controlled on-road driving event in drivers with both visual field loss and age-matched healthy controls (Charlton, 2008). The driving task involved several hazardous events including a car (target) exiting a driveway, stopping just short of the driver's lane. The timing and location of the hazardous events in the on-road drive was controlled as tightly as possible to match the simulated drive conditions using a system of road markers and use of interactive signals between the test vehicle and target vehicle. Inspection of the data in Figure 13.2 shows individual Case and Control participant profiles with a relatively tight coupling of the start point and rate of the deceleration in response to the unexpected incursion of a car at the side of the road next to the driver's lane. Case and Control participants reached their slowest approach speeds slightly earlier (within 25 m) in the simulator than the on-road task.

Klee et al. (1999) examined the validity of the fixed-base simulator at the University of Central Florida with respect to forward speed. Thirty drivers (aged 17–65 years) drove the same section of campus road on the simulator and in an instrumented vehicle. Longitudinal speed was recorded at 16 locations along the track.

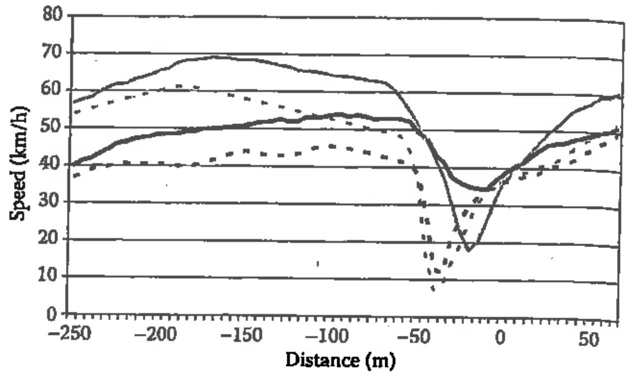

FIGURE 13.2 Individual Control (black) and vision-impaired Case (grey) participants' speed profiles for simulator (broken line) and controlled on-road (solid line) driving event involving a car exiting from a driveway next to the driver's lane at location 0 m.

Speeds in the two driving environments were similar at 10 locations, while at the remaining six locations participants drove at a slower speed in the simulator, leading the authors to conclude that absolute validity was not demonstrated. While the authors make no specific claims about relative validity, the results show that at the majority of data sites (10 out of 16) participants drove approximately 5–10 km/h slower in the simulator than on-road, and the speed differential was the same, implying relative validity. Bella (2005) provided further evidence for relative coupling of speed in a simulator and an on-road environment in a highway driving context. Thirty-five drivers (aged 24–45 years) completed a 12 km highway drive through a work zone from which speed recordings were obtained every five meters. A bilateral Z-test for non-matched samples was performed in order to estimate whether the driving speeds in the field were significantly different from those in the simulator. Consistent with Klee et al. (1999) findings, Bella (2005) reported that the simulator speeds tended to be lower than the speeds recorded in the field drive. However, the difference between the mean speeds in the simulator compared to the mean speeds in the field were not significantly different at each measurement site, demonstrating, at a minimum, interactive relative validity of the interactive static base simulator for assessing speed.

Yan, Abdel-Aty, Radwan, Wang and Chilakapati (2008) investigated the validity of the University of Central Florida's driving simulator for evaluating aspects of traffic safety at signalized intersections. A real-world intersection was replicated in a simulator, with eight intersection scenarios. Participants' speed (in four scenarios) and risky traffic behaviors (in four scenarios) were recorded. Absolute validity was demonstrated for the majority of intersection approach speed measures. Comparisons between participants' approach speed in the four real and simulated scenarios showed that all speed distributions followed normal distributions, there were no statistically significant differences in three of the four scenarios for mean approach speed (the fourth scenario approached significance at $p = 0.051$), and no differences in two of the four scenarios for speed variance. Additionally, Yan et al. (2008) found evidence for relative validity for selected measures of risky driving behavior. Comparisons of simulator driving behavior at two right-turn lanes—one with a real-world low rear-end crash record (two crashes) and one with a real-world high rear-end crash record (16 crashes)—revealed a higher number of risky behaviors (e.g., higher deceleration rate, faster speed at the stop line, closer following distance) in the lane with the higher real-world crash rate. That is, the same performance patterns observed in the real-world low and high risk intersections were upheld in the simulator.

Rather than using more conventional ways of measuring speed as an indicator of behavioral validity, Shinar and Ronen (2007) examined two speed-related measures: speed estimation (i.e., driver estimations of the vehicle's current speed) and speed production (i.e., adjustment of the vehicle's speed to achieve a predetermined speed). Sixteen participants (aged 24–30 years) completed an on-road and a simulated drive (in a full-cab simulator with STISIM software) with the speedometer shielded from view. During each drive, participants were required to estimate and produce seven different speeds (40, 50, 60, 70, 80, 90, and 100 km/h). The authors found that speed estimation performance in the simulator only differed significantly from on the real road for the speeds of 40 km/h (where higher speeds were estimated in the simulator) and 100 km/h (where lower speeds were estimated in the simulator). For speed production, produced speeds were consistently approximately 25 km/h faster in the simulator than on-road. This study shows absolute validity for most speeds for speed estimation, and relative validity for speed production.

13.3.2 Lateral Position

Another common measure for validation is the extent to which the lateral position of the vehicle is matched across simulator and real driving contexts. Blana and Golias (2002) investigated the validity of the fixed-base Leeds Advanced Driving Simulator for assessing vehicle lateral displacement (measured as the distance from the front tire on the passenger side of the vehicle to the white line at the edge of the road) on straight and curved sections of the road. One hundred on-road drivers were observed (with video cameras positioned on the ground and mounted on lamp posts) driving on one straight and two curved sections of a road. A further 100 participants (mean age of 36 years) completed a simulated version of the on-road drive. The results showed that for both straight and curved sections, the mean lateral displacement was significantly larger on the real road than in the simulator (suggesting that on-road drivers positioned their vehicles further from the road edge, closer to the centerline), while the standard deviation (SD) of lateral displacement was significantly smaller on the real road than in the simulator (suggesting that on-road drivers maintained a more consistent lane position). However, speed influenced the results. For straight sections, the mean and SD of lateral displacement only differed between the real and simulated roads when speeds were greater than 70 km/h. For curved sections, the mean lateral displacement differed significantly only when speeds were less than 60 km/h (speed did not affect results for the SD of lateral displacement on curved sections). Although the direction of the difference in lateral displacement between the on-road and simulated drives was consistent across all speeds (i.e., the on-road mean lateral displacement was always larger than the simulated drive, and the on-road SD of lateral displacement was always smaller), the amount of difference in lateral displacement varied depending on speed, and thus neither absolute nor relative validity was demonstrated.

Wade and Hammond (1998) also examined simulator validity for measuring lane deviation (measured as the vehicle's distance from the centerline). They assessed the driving performance of 26 participants (aged 18–34 years) on a real-world driving route and a comparable virtual driving route programmed on the University of Minnesota's Human Factors Research Laboratory wrap-around simulator. The mean deviation from the centerline was statistically significantly larger in the virtual than the on-road environment (suggesting that on-road drivers positioned their vehicles closer to the centerline, further from the road edge), thus absolute validity was not demonstrated. However, mean deviations on four different types of roads showed a similar pattern in the two environments (e.g., the road with the highest mean deviation in the real world was also the road with the highest mean deviation in the virtual environment), thus demonstrating relative validity. It is important to note, however, that data were collected only on straight road segments with maximum speed limits of 35 mph (approximately 56 km/h) and where drivers would have a constant speed. Whether speed would influence validity for measures of lateral position (as found in Blana & Golias, 2002) remains unknown. Wade and Hammond (1998) also observed that drivers who displayed an aggressive driving style on-road (e.g., hard acceleration, hard braking, and speeding) showed similar aggressive behaviors in the simulator. Furthermore, similar habitual behaviors were observed in the on-road and simulated setting; participants who steered away from the centerline when approaching a real stop sign also displayed this behavior in the simulator. Participants also displayed reflexive responses to deceleration (i.e., they flexed their arm and shoulder muscles) when braking in the simulator. Reflexive responses in real-world driving prevent drivers moving forward due to inertia when they brake. There was no inertia in the simulator, and hence reflexive responses were not required, yet participants performed this habitual behavior. It is clear that the simulator evoked many behaviors of drivers in the real world, suggesting the simulator was behaviorally valid. Overall, these studies suggest that simulators can show relative validity for measures of lateral position, but they are unlikely to show absolute validity; variation in lane position is likely to be greater in simulated than real-world environments, and drivers are likely to travel further from the centerline in the simulated world than in the real world.

13.3.3 Braking Responses

In some studies, braking responses offer an obvious measure for simulator validation, particularly for driving tasks where rapid and timely braking is critical for safety to avoid a crash. Hoffman, Lee, Brown and McGehee (2002) investigated the braking responses of drivers in an on-road and a simulator scenario to validate the Iowa Driving Simulator for braking. Sixteen participants (aged 25–55 years) were instructed to brake normally or hard to avoid a collision with a leading vehicle. The time headway (the time between the two vehicles when the lead vehicle began to break) was fixed at 1.7 seconds in the simulator. The time headway on the on-road course could not be fixed, but drivers had a mean time headway of 1.6 seconds, suggesting that the fixed simulator headway of 1.7 seconds was representative of real driving conditions. Performance was visually compared on several measures of brake onset, brake profile, and brake completion. The general pattern of results in the simulated and on-road courses was similar (most measures showed a difference between simulator and on-road performance of similar magnitude), and the driver's initial speed (30 or 60 mph) and the deceleration rate of the leading vehicle (0.15 or 0.40 g) affected braking responses similarly in the two driving environments (i.e., most means for simulator and on-road performance were in the same direction for each level of the independent variables), suggesting evidence of behavioral validity. However, it is difficult to confidently evaluate validity using only visual inspection to describe observed differences. In this study, simulator data were compared with data previously collected on-road; hence detailed data were available only for the simulated drives. Analyses were limited to graphed means for on-road data, and graphed means with error bars for simulator data. Although this enables some conclusions to be drawn regarding the simulator's validity, it is preferable to statistically analyze data, and the absence of statistical analyses should be considered a limitation of the study.

Regarding the effect of the braking instructions (normal or hard), drivers in Hoffman et al. (2002) study responded differently in the two environments. For the hard braking instruction, data were similar in the simulator and on the road. However, for the normal braking instruction, participants in the simulator responded in the same way as they did to the hard braking condition. Given that the simulator did not provide the same cues associated with normal braking, including haptic and vestibular feedback, participants may have found it difficult to implement the different degrees of braking required.

Braking and accelerator usage were also studied by McGehee et al. (2000) to evaluate the validity of the Iowa Driving Simulator for a crash-avoidance task. Participants completed a simulator drive that involved an intersection crash scenario, and an on-road drive that included a similar crash scenario. The on-road drive required participants to maintain a headway of two seconds behind a lead vehicle on approach to an intersection. Participants drove through the intersection a number of times while navigating through real-world traffic. The crash scenario in the on-road drive was constructed using a foam vehicle that replaced one of the real cars on the final lap and was propelled in front of the driver's vehicle. The time period from when the driver released the accelerator and put maximum pressure on the brake did not significantly differ between the simulator and on-road drives, nor did the point at which drivers first used the steering wheel to avoid crashing. However, the time to accelerator release was significantly greater in the on-road drive compared to the simulator drive. The findings suggest that absolute validity was upheld for the timing of the drivers' transition from accelerator to maximum braking but not for the initial response time for foot off accelerator. The authors suggested that this might be due to methodological differences between the conditions, with participants completing three laps of a course in the on-road drive and one lap of a track in the simulator drive. Drivers in the on-road course might therefore have been more familiar with the intersection event, as they would have already driven through the intersection two times before the crash scenario occurred.

13.3.4 Validity of Using Simulators for Assessing Road Safety Countermeasures in the Real World

The validity of simulators for assessing the effects of road design and traffic control devices has also been examined. Riemersma, van der Horst, Hoekstra, Alink and Otten (1990) examined the validity of the Daimler-Benz advanced driving simulator for assessing speed reduction methods. A real-world study investigated the effect of infrastructure changes on drivers' speed when entering a Dutch village. The infrastructure changes to reduce speed included a median strip, colored asphalt, and a portal gate. Comparisons of vehicle speeds prior to and following installation of these infrastructure changes showed that the average speed of drivers upon entering the village decreased following installation, and faster drivers reduced their speed more than average-speed drivers. Riemersma et al. (1990) replicated this scenario in their simulator. Twenty-four male participants (aged 20–35 years) drove for approximately one hour around a course that required them to pass through the entrance to the village 12 times. Each approach consisted of a different combination of infrastructure changes. Comparing results from the real-world and simulated studies showed that on approach to the entrance of the village (i.e., 400 m from the village entrance), participants drove at a faster speed in the simulator than drivers in the real-world environment. However, simulator participants then reduced their speed to a greater extent than drivers in the real world, so that their speed at the entrance point to the village was slightly slower than drivers in the real world. The infrastructure changes produced larger speed reductions in the simulated than the real-world environment, with the mean entrance speed decreasing 25.7% in the simulator (from 78.3 km/h to 58.2 km/h) and 8.6% in the real world (from 72.4 km/h to 66.2 km/h), thus demonstrating relative but not absolute validity. In addition, the difference in entrance speed between on-road fast and average-speed drivers (approximately 10 km/h) was similar to the difference in entrance speed of simulator drivers when instructed to drive "as quickly as the conditions would allow" versus "in a relaxed and unhurried manner" (p. 418). The authors concluded that the simulator was an effective tool for evaluating speed reduction methods. Similarly, as described above, Godley et al. (2002) demonstrated the relative validity of the MUARC simulator for examining the effect of rumble strips as a speed-reduction method. In contrast, McAvoy, Schattler and Datta (2007) found that their Doron Precision Systems, Inc. AMOS II simulator was not valid for evaluating night traffic speed control devices at construction zones (see also Wood & Chaparro, this book, chap. 28). Six real-world work zones were incorporated into a driving simulator scenario. Radar guns measured the speed of motorists travelling through the real-world work zone sites at night. Mean speeds in these field sites differed according to the presence or absence of steady-burn warning lights mounted on drums; however, the warning lights did not significantly affect mean speeds in the simulator. The simulator failed to meet standards for absolute or relative validity, perhaps because it failed to induce the perception of risk associated with driving through real-world work zones at night.

13.3.5 Validity of Using Simulators for Assessing Complex Driving Behaviors in the Real World

The validity of simulators for assessing more complex behavior, such as performance during divided attention tasks, has also been examined. Reed and Green (1999) investigated the validity of the University of Michigan Transportation Research Institute driving simulator for assessing decrements in driving performance during a manual car phone task. Younger (20–30 years) and older (>60 years) participants were instructed to dial a manual car phone while driving in an on-road setting and a simulator setting. Results showed that speed control (i.e., SD of speed and SD of throttle position) was similar in the two driving environments but lane keeping (i.e., mean lateral speed and SD of steering wheel position) varied more in the simulator. For seven of ten variables measuring lane position, speed, steering wheel angle, and throttle position, medium to high correlations were found between on-road and simulated performance ($r = 0.43$ to $r = 0.76$). For both age groups in both driving environments, the phone task resulted in decrements in speed control and lane keeping performance; however decrements were larger in the simulator. In addition, older drivers displayed greater decrements in driving performance than younger participants. Admittedly this is a simplified summary of their results (simulator fidelity was a further independent variable), but overall, absolute validity was demonstrated for speed control, while relative validity was shown for assessing the effects of dialing a phone and age on driving performance. Unfortunately, results for only four of the ten measured variables were presented in Reed and Green (1999) report (following analysis of all ten variables, one variable was selected to illustrate the full results for lane position, speed, steering wheel angle, and throttle position). There are few other validation studies investigating complex behavior, suggesting a need for further research in this area.

13.3.6 Validity of Using Simulators for Assessing Driving Behavior of Specific Driver Groups

Driving simulator validation studies typically involve young to middle-aged drivers. Increasingly, attention has been focused on studying driving patterns of specific sub-groups of drivers at high risk, based on variables such as age, medical conditions, and impairments. For vulnerable and potentially high risk drivers, the safety afforded by the simulator as a research tool is highly valuable. However, the usefulness of this research for understanding the impact of aging and impairments on driving behaviors hinges on the assumption that performance in the simulator reflects real-world driving. The accuracy of this assumption is especially critical if simulator performance outcomes are to be used for decisions about fitness to drive (e.g., see Ball & Ackerman, this book, chap. 25). It is possible that simulated driving conditions may differentially affect drivers with certain kinds of conditions—such as visual field loss—in a way that is not evident in a naturalistic driving setting; hence the importance of demonstrating the validity of the simulator, not just for specific driving tasks and outcome measures, but potentially for specific groups of drivers.

Hakamies-Blomqvist, Östlund, Henriksson and Fildes (2001) investigated the validity of the VTI simulator as a tool to study driving behavior of older drivers, taking into account participants' histories of accidents and incidents, and the difficulty experienced driving in stressful situations. One issue with assessing simulator validity for older drivers is variation in driver performance. Older drivers can be expected to exhibit greater variation in driving performance than younger drivers due to age-related declines in cognition, psychomotor function, vision, and health. Thirty-five drivers (aged 66–80 years) completed an on-road drive and a corresponding simulator drive. Participants drove in a similar way in the two settings, although in the simulator the mean speed was lower and there were greater variations in speed, lateral distance, and steering wheel movements. Participants also used the brake more frequently and with greater force in the simulator. Relative validity was obtained for operative measures such as the steering wheel angle and lane position. Results also showed that validity was strongest for drivers who reported the fewest number of driving crashes and incidents, and who reported no difficulty driving in stressful situations. These results suggest that the less competent the driver, the less valid the simulator is as a screening tool (see also, in this book, chap. 50 by Brouwer et al., and chap. 26 by Trick & Caird). This has implications for testing cohorts with poor driving ability (e.g., some drivers recovering from a stroke).

Recent research with the STISIM 400 simulator found that the on-road driving behavior of eight older adults (aged 67–81 years) was highly correlated ($r = 0.74$) with driving behavior in the simulator (Bédard et al., 2010). Although this analysis lacks power with a small sample size, there was a significant association ($p = 0.035$) between demerit points recorded during the on-road drive and a simulated replica. Lee and colleagues also conducted a number of studies involving older adults using the STISIM driving simulator. Lee, Cameron, et al. (2003) found that behavioral measures recorded from a driving simulator were highly correlated with measures recorded during an on-road drive, indicating that the STISIM simulator was a valid tool for assessing older driver behavior.

In subsequent research, Lee and colleagues examined the correspondence between simulator driving tasks and self-reported traffic violations and crash history (Lee & Lee, 2005; Lee, Lee, Cameron, & Li-Tsang, 2003). In a retrospective study, Lee, Lee, Cameron and Li-Tsang (2003) assessed older adults' driving simulator performance and crash history. One hundred and twenty-nine participants (aged 60–88 years) participated in a 45-minute driving simulator session, which was monitored by an assessor. Driving measures recorded during the simulated drive included compliance to traffic signs, driving speed, correct use of the indicator, ability to drive through a T intersection, working memory, speed limit compliance, rapid decision-making ability, functional reaction time, ability to perform simultaneous tasks, and knowledge of—and compliance with—traffic regulations. Prior to the drive, participants were interviewed about their crash history, with over 60% reporting involvement in at least one crash in the previous year. Logistic regression analysis showed that working memory, speed limit compliance, and rapid decision-making ability were significant predictors of reporting a prior crash. All associations were negative, such that participants were more likely to report prior crash involvement if they had poorer working memory (assessed by reporting back five street signs after ten minutes), poorer speed limit compliance, and poorer ability to make rapid decisions and judgments to avoid dangerous traffic situations in the simulated drive.

Following the retrospective study, Lee and Lee (2005) conducted a prospective study to determine whether simulator driving behavior could identify potentially at-risk older drivers. The same 129 drivers from Lee, Lee, Cameron and Li-Tsang (2003) completed a 45-minute simulated driving task, an interview, and a feedback session. Driving measures were recorded as per Lee, Lee, Cameron and Li-Tsang (2003). Participants were asked about their prior vehicle crashes and traffic violations, and gave permission for the researchers to access their driving records for the following three years. All participants received at least one traffic violation during the three-year period. As expected, driving task performance was negatively associated with age. The results also revealed that working memory ability and appropriate use of the indicator were negatively associated with the number of traffic violations over the three years. The authors attributed these findings to an age-related decline in psychomotor coordination and reaction time, and concluded that the simulator could identify older drivers at risk of future traffic violations.

Regarding the validity of simulators for evaluating the effects of training-specific cohorts, research with novice drivers shows that training effects measured on simulators correspond with effects measured on-road. Fisher, Pradhan, Pollatsek and Knodler (2007) report two studies (an on-road and a simulated study) where the same PC-based training program was used to train novice drivers (aged 18–21 years) to identify potential hazards. A mobile eye tracker collected eye movement data, so that participants' scanning behavior could be examined. In the on-road study, during a post-training on-road drive, 12 participants who completed the training program scanned critical areas significantly more often than 12 control (untrained) participants (mean scanning time of 64.4% and 37.3% respectively). For the simulated study, post-training performance was tested on the University of Massachusetts at Amherst advanced fixed-base driving simulator. During the post-training simulated drive, six trained participants scanned critical areas of the road significantly more often than six untrained participants (mean scanning time of 77.4% and 40.0% respectively). Although only a simplified version of the results are reported here (scanning in different types of scenarios was a further variable examined), the results demonstrate a degree of correspondence between training effects measured on the simulator and those measured on-road. The magnitude of training effects measured in each drive did not statistically significantly differ, although training effects seemed larger when measured on the simulator (training improved scanning by a mean of 37.4% on the simulator versus a mean of 27.1% on-road). This may be due to the limitations of the study such as small sample sizes, and having different participants complete the on-road and simulated drives.

13.3.7 Crash and Infringement History

Validation methods based on comparisons of direct measures of driving behaviors in simulated driving against on-road driving behavior are arguably more robust than methods based on comparisons of driving behaviors in simulated driving with driving behavior questionnaires and crash history data (Anstey, Wood, Lord, & Walker, 2005). Crash history data and participant reports of traffic violations are not always adequate representations of individual driving skills. External factors can contribute to driving accidents, and people may not report all previous traffic violations (Lew et al., 2005). Despite their less robust design, the results of studies assessing simulator validity using comparisons of simulator behavior and behavior reported in questionnaires are encouraging (Lew et al., 2005; Reimer, D'Ambrosio, Coughlin, Kafrissen, & Biederman, 2006). Reimer et al. (2006) found high correlations between behaviors observed in a driving simulator and participant self-reports of driving behavior and accident history (valid measures included speeding, velocity, passing, traffic weaving, pause time at stop signs, and accidents). Furthermore, participants diagnosed with Attention-Deficit Hyperactivity Disorder (ADHD; who are more likely to report involvement in previous crashes than individuals without ADHD) were more likely than control participants to be involved in a crash on the simulated drive. However, the authors did acknowledge one issue with using questionnaires—there was some discrepancy between some questionnaire items and the corresponding simulator measure.

To ensure simulators are useful tools for assessment, researchers have examined the ability of simulators to predict future driving behavior and to distinguish between safe and unsafe drivers. Lew et al. (2005) conducted a prospective study comparing participants' simulator performance (recorded by the simulator and an observer) and on-road performance (recorded by an observer) with their on-road driving performance 10 months later. Both simulator- and observer-recorded measures of simulator driving performance predicted on-road driving performance 10 months following the simulated drive. However, simulator-recorded measures were more sensitive and accurate than observational measures. The predictive efficiency of the simulator-recorded measures was 82%, with sensitivity of 100% (i.e., all participants who failed the simulator drive were classified as failing the on-road drive 10 months later) and specificity of 71% (i.e., 71% of participants who passed the simulator drive were classified as passing the on-road drive 10 months later). On-road performance was not correlated with on-road performance 10 months later. These results suggest that driving simulator measures can be better predictors of future driving performance than data obtained from on-road drives. Other studies have similarly found that simulator driving measures can identify drivers at risk of future traffic violations (e.g., Lee & Lee, 2005) and distinguish between drivers of differing abilities (e.g., Lee, Lee, Cameron, and Li-Tsang 2003; Patomella, Tham, & Kottorp, 2006.)

13.3.8 Physiological Measures

Physiological responses have also been used to examine behavioral simulator validity. As noted above, while this is a less direct way of measuring driving behavior, physiological responses can be used to determine the participant's awareness level in the simulated environment, and thus provide some useful insights into behavior across simulated and real-world driving contexts. Valid simulations should induce similar physiological responses as in the real world. Watts and Quimby (1979) validated the Transport and Road Research Laboratory simulator using a measure of skin conductance. Participants assessed risks during an on-road drive and while viewing a 30-minute film of the same drive on the simulator. The film showed 45 hazardous scenarios from the driver's perspective and participants indicated the amount of perceived danger on a 10-point scale. During the film, participants' levels of skin conductance were measured. Risk assessment ratings were similar in the two environments and perceived risk rankings were highly associated with levels of skin conductance ($r = 0.78$). In another validation study with physiological measures, Bédard (2008) reported on the measurement of heart and respiration rates of participants exposed to threatening situations during a drive on the STISIM 400 simulator. The threatening situations included two green traffic lights that changed to amber as the driver approached the intersection, and a car that pulled out from the shoulder of the road in front of the driver. Each threatening situation increased participants' mean heart rate and two of the three situations increased participants' mean respiration rate, indicating that the simulator induced a high degree of presence. A third study included measures of both skin conductance and heart rate during an on-road drive and a simulated drive on the DriveSafety DS-600c motion-based simulator (Slick, Evans, Kim, & Steele, 2006). During four specific driving tasks (right and left turns at stop signs and traffic lights), Slick et al. (2006) found that skin conductance and heart rate did not significantly differ between the simulated and on-road drives. Similar results were found for mean changes in skin conductance and heart rate during the entire drive for both males and females, with the exception that female participants' mean change in heart rate was significantly smaller in the simulator drive than the on-road drive. Overall, these studies suggest that physiological measures generally demonstrate the validity of simulators.

13.3.9 Ecological Validity

As an alternative to directly measuring and statistically comparing simulator and on-road behavior, some studies have examined simulator validity by measuring only simulator performance (e.g., the speed at which male and female participants drive in a simulator), and comparing this behavior with pre-established on-road behavior (e.g., the general trend for men to drive faster than women on-road). This examines a form of *ecological validity*, defined here as the degree to which simulator behavior reflects real-life on-road behavior patterns displayed over extended periods of time (as opposed to examining simulator and on-road performance at a single point on a drive; Lew et al., 2005). In this way, simulators can be validated for general patterns of behavior, without directly measuring on-road performance as part of the study. This method is a less direct and robust means of establishing validity than those previously discussed. Nonetheless, it can be a useful first step in assessing a simulator's validity, enabling conclusions regarding the overall pattern of behavior in the simulator compared with behavior in the real world. There is a vast amount of research on ecological validity and a full account of this literature is beyond the scope of this chapter. The following studies illustrate the general concept.

The effect of age on driving is well established in the real world. For example, younger drivers generally react more quickly than older drivers, and they tend to drive at faster speeds. Studies conducted on the STISIM 400 driving simulator have found similar age effects. Toxopeus (2007) examined braking responses to a stop sign in younger ($M = 22.1$ years) and older ($M = 69.1$ years) participants. A stop sign was displayed at the end of a 45-minute simulator highway drive and participants were required to stop as quickly as possible. Younger participants ($M = 0.92$ seconds) pressed the brake pedal significantly more quickly than older participants ($M = 1.01$ seconds), demonstrating behavioral validity for age effects on braking responses. Younger drivers ($M = 1.27$ seconds) also took less time to press the horn in response to a visual cue than older drivers ($M = 1.66$ seconds). Furthermore, when asked to maintain a speed of 55 mph, younger adults ($M = 56.4$ mph) drove faster than older adults ($M = 53.9$ mph), and there was a trend towards males ($M = 56.1$ mph) driving faster than females ($M = 54.7$ mph; $p = 0.055$).

Also using the STISIM driving simulator, Lee, Lee and Cameron (2003) examined visual attention ability in older adults. They had 129 adults aged over 60 complete a 45-minute simulated drive. A visual stimulus was presented 14 times during the drive, and participants were instructed to engage the turn indicator as soon as they saw the stimulus appear. Results showed a positive association between age and reaction time, with older seniors taking more time to respond than younger seniors. As a decline in visual attention was consistent with the aging literature, Lee, Lee and Cameron (2003) concluded that the STISIM simulator could be used to investigate visual attention ability in older drivers.

Demographic effects were also found by Yan and colleagues (2008) using the University of Central Florida driving simulator. In the simulated intersection scenarios, males drove faster than females, and participants in the 20–24 years age group drove faster than participants in older age groups. The consistency of these findings across simulators and studies suggests simulators do reflect similar behavior patterns as those seen in real-world driving.

Another means of evaluating simulator validity without on-road performance measurement is to relate simulator behavior with cognitive measures known to predict on-road driving performance (see also Ball & Ackerman, this book, chap. 25). For example, performance on the Useful Field of View® test (UFOV) is a good predictor of performance (pass/fail) in on-road driving evaluations (Myers, Ball, Kalina, Roth, & Goode, 2000), and is also associated with the frequency of on-road at-fault crashes (Clay, 2005). Bédard et al. (2010) reported two studies demonstrating strong correlations between simulator and cognitive measures such as UFOV. Participants (aged 18–83 years) drove a simulated road test route (i.e., a simulated version of the Manitoba road test, which is used for driver licensing) through their local city. The simulator recorded their driving errors and an observer recorded their demerit points. Performance on two cognitive tests—UFOV and Trail Making Test Part A (Trails A)—was significantly correlated with simulator-recorded driving errors and observer-recorded demerit points (r's ranged from .57 to .74). Wald and Liu (2001) also found that simulator driving performance was correlated with the Trails A scores. Performance on the Attention Network Test (a cognitive test that has good concurrent validity with UFOV; Weaver, Bédard, McAuliffe, & Parkkari, 2009) has also shown significant associations with simulator driving performance (Mullen, Chattha, Weaver, & Bédard, 2008; Weaver et al., 2009). However, validity assessments should not assume that all relationships are linear (the reader will recall our earlier discussion of functional validity). Weaver et al. (2009) considered both linear and quadratic associations between simulator driving performance and UFOV scores, and found the latter relationship was stronger (linear $r = 0.73$, quadratic $R = 0.79$). Figure 13.3 similarly demonstrates the importance of considering alternative associations.

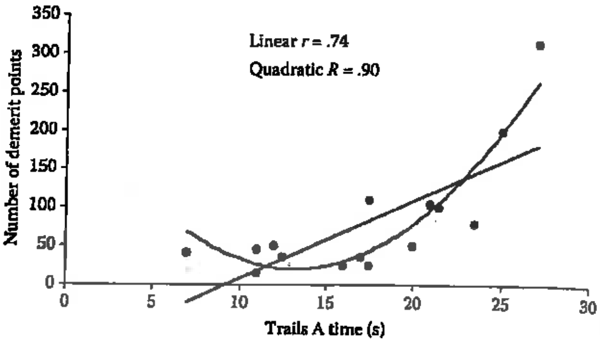

FIGURE 13.3 Relationships between simulator-recorded demerit points and Trails A.

13.3.10 Summary

In summary, there is sufficient evidence to suggest that simulators provide a valid tool for assessing a variety of driving performance measures such as speed, lateral position, and risky traffic behaviors. They also appear valid for assessing the effects of divided attention on driving performance. Furthermore, measures from simulator drives can identify older drivers at risk of future traffic violations, and simulators are sensitive to age-related changes in driving performance and cognition. Simulators have been validated for assessing age effects on speed, braking responses, reaction time, and some operative measures (e.g., vehicle position on the road), although the evidence is less convincing for older drivers who experience difficulties with driving. Regarding braking responses, validity has been demonstrated for brake onset but not for braking force (i.e., braking normally versus hard). Although simulators appear valid for assessing the effects of some traffic control methods (e.g., rumble strips), they are not valid for all devices (e.g., lights at construction zones). The majority of validated measures show relative validity but fail to meet requirements for absolute validity.

The driving ability of participants can compromise simulator validity, such that simulators may be less valid for poorer drivers. In addition, variability in driving performance measures has a tendency to be greater in simulator than on-road drives. Hence researchers, driving evaluators, and other simulator users should remain aware that simulators do not always provide an accurate picture of on-road driving behavior.

13.4 Future Research Directions

Although Yan et al. (2008) established absolute validity for mean intersection approach speed, speed variability was comparable at just two intersection approaches with lower speed limits (45 mph). At two intersection approaches with higher speed limits (50 mph), simulator speeds had larger variability than on-road speeds. These results suggest that future research should investigate the variability of data in addition to measures of central tendency because these factors can affect validity conclusions.

The usefulness of simulators as a tool for assessment and training depends on simulator validation for the cohort being assessed or trained. Simulators offer a safe alternative to on-road training for cohorts who could be hazardous to themselves or other road users in real-world environments. For example, fatigue effects and the effects of medications on driving performance can be safely monitored in a simulator environment. Philip and colleagues (2005) were interested in the effects of sleep deprivation on driving behavior in an on-road and a simulated driving environment (Philip et al., 2005). Twelve men aged 19–24 years were divided into a regular sleep condition of eight hours per night and a restricted sleep condition consisting of two hours of sleep. Measurements included self-rated fatigue, simple reaction time, and the number of inappropriate line crossings. Participants' self-rated sleep scores, reaction times, and inappropriate line crossings were significantly higher in the simulator than on the road, but the effect of sleep deprivation was comparable in the simulated and on-road drives. In contrast, Volkerts and colleagues (1992) found the TS2 driving simulator was a less sensitive measure of sedative drug-induced impairment than an on-road drive (Volkerts, van Laar, van Willigenburg, Plomp, & Maes, 1992). Oxazepam (50 mg) and lormetazepam (1 mg) impaired on-road driving performance, but did not impair performance in the simulator. These findings highlight the need for further research to establish the validity of driving simulators for measuring the effects of medications.

While there are some validation studies examining older drivers, there is a need for future research to examine cohorts with specific medical conditions (e.g., stroke, Alzheimer's disease, acquired brain injury). Assessment of driving skills on a simulator may influence decisions concerning whether a driver's license may be retained or revoked, hence the importance of establishing validity with these cohorts (see also Brouwer et al., this book, chap. 50). Lee et al. (2007) examined the simulator and on-road driving performance of participants with and without Parkinson's disease (PD). Participants with PD had significantly poorer driving performance than healthy participants in both the on-road and simulator environments. However, simulator performance explained 68% of the variability in on-road performance for healthy participants but only explained 39% of the variability for participants with PD. These results suggest that the simulator may be less valid for participants with PD, and highlight the need for future research to examine the effect of other medical conditions on validity.

Simulator discomfort (SD) remains an unresolved issue for all cohorts, even those for whom validity has been established (see also Stoner, Fisher, & Mollenhauer, this book, chap. 14). SD refers to a range of symptoms involving stomach, oculomotor, and orientation disturbances (e.g., nausea, headache, dizziness; Kennedy, Lane, Berbaum, & Lilienthal, 1993); one suggested explanation, but not the only one, is that it results from the conflict between current sensory information (e.g., visual information suggesting the individual is moving and vestibular information suggesting the individual is stationary) and stored sensory expectations based on past experience (Reason & Brand, 1975). Symptoms can be sufficiently severe that participants are unable to complete the simulator drive. Alternatively, affected participants may be able to complete the drive but the effect of SD on their driving performance is unknown. Although SD can potentially affect any simulator user, studies have found that variations among individuals, simulator set-ups, and simulator tasks influence susceptibility. For example, Park, Cook, Rosenthal, Fiorentino and Allen (2006) found a higher incidence of SD in a scenario involving high speed driving in a complex environment (city streets with high traffic density and multiple turning maneuvers) than in slower, less complex scenarios with fewer maneuvers. In a sample of older drivers (284 drivers aged 60–99) referred for driving evaluation, Freund and Green (2006) found almost 11% reported experiencing SD, with reports significantly greater among females than males. More than half (57%) of the participants who reported experiencing SD were unable to complete the drive. SD prevents otherwise valid simulators from being valid for all individuals and scenarios. Therefore, to improve simulator validity, future research should investigate methods to reduce or eliminate the occurrence of SD.

Although absolute and relative validity are the most common types of behavioral validity examined, and hence have been discussed throughout this chapter, Kantowitz (2001) offers another view of behavioral validity (see also, Kantowitz, this book, chap. 3). While acknowledging that the distinction between absolute and relative validity is valuable, he states:

> . . . I prefer to think of simulator validity more in terms of regression. How well does the simulator predict an outcome on the road? This allows for outcomes that are not absolute but still are better than relative validity . . . . Regression analysis also offers a metric that explains how well simulators predict reality so that different users can make their own judgments about the sufficiency of the fit for their own design purposes (Kantowitz, 2001, p. 51).*

We encourage researchers to consider this approach when examining simulator validity. Establishing whether a simulator exactly replicates on-road performance (absolute validity) is less important than determining *how well* a simulator predicts on-road performance; and this approach would provide a numerical measure. Regression analysis would also allow for examination of a range of relationships between simulator and on-road performance (e.g., linear, quadratic; as previously mentioned in our discussion of functional validity).

13.5 Conclusion

Many studies have investigated the validity of simulators for examining driving behavior. Most studies support the use of simulators, finding that driving behavior in simulators approximates (relative validity), but does not exactly replicate (absolute validity), on-road driving behavior. Simulator driving performance shows medium to strong correlations with a range of on-road driving performance measures, in addition to cognitive and physiological measures. This is sufficient for the majority of research, training, and assessment purposes for which simulators are used. However, a simulator shown to be valid in most settings is not guaranteed to be valid in the next setting. There can be critical differences (e.g., the environment, equipment, or protocol used), and when safety recommendations are based on simulator research, simulators almost necessarily require validation because the settings and scenarios are always changing.

Following simulator validation, adherence to the validated protocol is necessary to ensure measurements remain valid. Deviations from the protocol (e.g., altering the simulator set-up or changing performance measurements) will require validity reassessment. Validity has also generally been established with healthy cohorts. When using simulators with cohorts that deviate from this standard, validity will need to be established prior to evaluating their driving performance. In addition, where absolute values are required, on-road measures will generally be necessary. As simulators continue to advance technologically, their validity is only likely to broaden in terms of performance measures and cohorts, ensuring that simulators increase their role as a tool for driving performance measurements.

Key Points

- Simulators need to be validated to ensure the results of simulator studies generalize to driving in the real world. Behavioral validity refers to the extent to which the simulator induces the same driving behaviors that occur when driving in the real world. Behavioral validity has been further defined in terms of absolute and relative validity. While absolute validity requires that the two driving environments produce the same numerical values, relative validity is established when the differences between the two environments are in the same direction, and of the same or similar magnitude. Absolute validity is rarely established in driving simulator studies, but establishing relative validity is both necessary and sufficient for simulators to be useful research tools in most, though not all, cases.
- Methodological considerations for validation studies include factors such as the nature of the research question, task conditions, and dependent measures. Participant characteristics, simulator equipment, and how driving behaviors are measured in the simulator and real world can also affect validity. Statistical techniques used to establish validity include analysis of variance, correlation, and descriptive analyses of observed measures (e.g., speed profiles) of behavior measured on-road and during simulated drives.
- Simulators appear to provide a valid tool for assessing a variety of driving performance measures including speed, lateral position, brake onset, divided attention, and risky traffic behaviors. The majority of validated measures show relative validity but fail to meet requirements for absolute validity. Simulators do not appear valid for measures of braking force or for assessing the effects of some traffic control methods. Simulators can also induce similar physiological responses to those expected in the real world. Simulators appear sensitive to age-related changes in real-world driving performance and cognition, and can identify older drivers at risk of future traffic violations.
- Future research should investigate the variability of data in addition to measures of central tendency and methods to reduce the occurrence of simulator discomfort, because these factors can affect validity conclusions. There is also a need to establish the reliability of using driving simulators with specific cohorts, such as drivers with specific medical conditions; research to date has generally established validity with healthy cohorts. Regression modeling offers a useful approach to exploring the question of validation and the extent to which variables measured in the simulator predict real-world driving.
- Most studies support the use of simulators, finding that simulator driving behavior approximates (relative validity), but does not exactly replicate (absolute validity), on-road driving behavior. This is sufficient for the majority of research, training, and assessment purposes for which simulators are used. However, simulator users should remain aware that simulators do not always provide an accurate picture of on-road driving behavior, and validity should not be assumed; each simulator set-up should be validated for its ability to measure the driving behavior of the cohort for which it is to be used. Furthermore, where absolute values are required, on-road measures will generally be necessary.

Keywords: Behavioral Validity, Dependent Measures, Real-world Driving Outcomes, Simulator Validity

Acknowledgments

Nadia Mullen acknowledges the support of a grant from the Natural Sciences and Engineering Research Council of Canada. Judith Charlton and Anna Devlin acknowledge the support of a grant from the Australian Research Council (Project ID LP0560272) and industry partners: GM Holden, VicRoads, Road Safety Trust (New Zealand), the Swedish Road Administration, and Zeiss. Michel Bédard is a Canada Research Chair in Aging and Health (http://www.chairs.gc.ca); he acknowledges the support of the Canada Research Chair Program. This research was also supported in part by grants from the Canadian Institutes of Health Research, the Canada Foundation for Innovation, AUTO21: Network of Centres of Excellence, and the Ontario Neurotrauma Foundation.

Glossary

Absolute validity: A type of behavioral validity; established when the simulated and on-road driving environments produce the same numerical values.

Behavioral validity: The extent to which a simulator elicits the same driving behavior that occurs when driving on real roads.

Ecological validity: The degree to which simulator behavior reflects real-life on-road behavior patterns displayed over extended periods of time.

External validity: The extent to which driving behavior in a simulator generalizes to real driving behavior.

Functional validity: A type of behavioral validity; established when the simulated and on-road driving environments produce numerical values that are associated with a mathematical function (linear or non-linear).

Interactive validity: A type of behavioral validity; established when the simulated and on-road driving environments produce the same numerical values (interactive absolute validity), or values similar in magnitude and in the same direction (interactive relative validity), over time.

Physical validity: The extent to which the physical components of a simulator vehicle correspond to on-road vehicles (includes consideration of the simulator layout, visual displays, and dynamic characteristics).

Relative validity: A type of behavioral validity; established when the simulated and on-road driving environments produce numerical values that are not identical, but are of similar magnitude and in the same direction.

Key Readings

Blaauw, G. J. (1982). Driving experience and task demands in simulator and instrumented car: A validation study. *Human Factors, 24*, 473–486.

Blana, E. (1996). *Driving simulator validation studies: A literature review.* Institute of Transport Studies, University of Leeds, UK.

Godley, S. T., Triggs, T. J., & Fildes, B. N. (2002). Driving simulator validation for speed research. *Accident Analysis & Prevention, 34*, 589–600.

Kaptein, N. A., Theeuwes, J., & van der Horst, R. (1996). Driving simulator validity: Some considerations. *Transportation Research Record, 1550*, 30–36.

Törnros, J. (1998). Driving behavior in a real and a simulated road tunnel—A validation study. *Accident Analysis & Prevention, 30*, 497–503.

References

Anstey, K. J., Wood, J., Lord, S., & Walker, J. G. (2005). Cognitive, sensory and physical factors enabling driving safety in older adults. *Clinical Psychology Reviews, 25*, 45–56.

Bédard, M. (2008, September). *Driving simulators: Are they useful tools?* Paper presented at the STISIM users group meeting, Québec City, Canada.

Bédard, M., Parkkari, M., Weaver, B., Riendeau, J., & Dahlquist, M. (2010). Assessment of driving performance using a simulator protocol: Validity and reproducibility. *American Journal of Occupational Therapy, 64*, 336–340.

Bella, F. (2005). Validation of a driving simulator for work zone design. *Transportation Research Record, 1937*, 136–144.

Bella, F. (2008). Driving simulator for speed research on two-lane rural roads. *Accident Analysis & Prevention, 40*, 1078–1087.

Bittner, A. C., Simsek, O., Levison, W. H., & Campbell, J. L. (2002). On-road versus simulator data in driver model development: Driver performance model experience. *Transportation Research Record, 1803*, 38–44.

Blaauw, G. J. (1982). Driving experience and task demands in simulator and instrumented car: A validation study. *Human Factors, 24*, 473–486.

Blana, E. (1996). *Driving simulator validation studies: A literature review.* Institute of Transport Studies, University of Leeds, UK.

Blana, E., & Golias, J. (2002). Differences between vehicle lateral displacement on the road and in a fixed-base simulator. *Human Factors, 44*, 303–313.

Charlton, J. L., Fildes, B., Oxley, J., Keeffe, J., Odell, M., Verdoon, A., et al. (2008, September). *Driving performance and visual search associated with visual field loss.* Paper presented at the International Conference on Traffic and Transport Psychology, Washington, DC.

Clay, O. J., Wadley, V. G., Edwards, J. D., Roth, D. L., Roenker, D. L., & Ball, K. K. (2005). Cumulative meta-analysis of the relationship between useful field of view and driving performance in older adults: Current and future implications. *Optometry & Vision Science, 82*, 724–731.

Fisher, D. L., Pradhan, A. K., Pollatsek, A., & Knodler, M. A., Jr. (2007). Empirical evaluation of hazard anticipation behaviors in the field and on driving simulator using eye tracker. *Transportation Research Record, 2018*, 80–86.

Freund, B., & Green, T. R. (2006). Simulator sickness amongst older drivers with and without dementia. *Advances in Transportation Studies*, Special Issue, 71–74.

Godley, S. T., Triggs, T. J., & Fildes, B. N. (2002). Driving simulator validation for speed research. *Accident Analysis & Prevention, 34*, 589–600.

Hakamies-Blomqvist, L., Östlund, J., Henriksson, P., & Heikkinen, S. (2001). *Elderly car drivers in a simulator—A validation study.* Linköping, Sweden: Swedish National Road and Transport Research Institute.

Hoffman, J. D., Lee, J. D., Brown, T. L., & McGehee, D. V. (2002). Comparison of driver braking responses in a high-fidelity simulator and on a test track. *Transportation Research Record, 2*, 59–65.

Hoskins, A. H., & El-Gindy, M. (2006). Technical report: Literature survey on driving simulator validation studies. *International Journal of Heavy Vehicle Systems, 13*, 241–252.

Kantowitz, B. H. (2001). Using microworlds to design intelligent interfaces that minimize driver distraction. *Proceedings of the international driving symposium on human factors in driver assessment, training and vehicle design, 2001* (pp. 42–57).

Kaptein, N. A., Theeuwes, J., & van der Horst, R. (1996). Driving simulator validity: Some considerations. *Transportation Research Record, 1550*, 30–36.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *International Journal of Aviation Psychology, 3*, 203–220.

Klee, H., Bauer, C., Radwan, E., & Al-Deek, H. (1999). Preliminary validation of driving simulator based on forward speed. *Transportation Research Record, 1689*, 33–39.

Lee, H. C., Cameron, D., & Lee, A. H. (2003). Assessing the driving performance of older adult drivers: On-road versus simulated driving. *Accident Analysis & Prevention, 35*, 797–803.

Lee, H., Falkmer, T., Rosenwax, L., Cordell, R., Granger, A., Vieira, B., et al. (2007). Validity of driving simulator in assessing drivers with Parkinson's disease. *Advances in Transportation Studies*, Special Issue, 81–90.

Lee, H. C., & Lee, A. H. (2005). Identifying older drivers at risk of traffic violations by using a driving simulator: A 3-year longitudinal study. *American Journal of Occupational Therapy, 59*, 97–100.

Lee, H. C., Lee, A. H., & Cameron, D. (2003). Validation of a driving simulator by measuring the visual attention skill of older adult drivers. *American Journal of Occupational Therapy, 57*, 324–328.

Lee, H. C., Lee, A. H., Cameron, D., & Li-Tsang, C. (2003). Using a driving simulator to identify older drivers at inflated risk of motor vehicle crashes. *Journal of Safety Research Senior Transportation Safety and Mobility, 34*, 453–459.

Lew, H. L., Poole, J. H., Lee, E. H., Jaffe, D. L., Huang, H. C., & Brodd, E. (2005). Predictive validity of driving-simulator assessments following traumatic brain injury: A preliminary study. *Brain Injury, 19*, 177–188.

McAvoy, D. S., Schattler, K. L., & Datta, T. K. (2007). Driving simulator validation for nighttime construction work zone devices. *Transportation Research Record, 2015*, 55–63.

McGehee, D. V., Mazzae, E. N., & Baldwin, G. H. S. (2000). *Driver reaction time in crash avoidance research: Validation of a driving simulator study on a test track.* Paper presented at the meeting of the Human Factors and Ergonomics Association/International Ergonomics Association, Santa Monica, CA.

Mullen, N. W., Chattha, H. K., Weaver, B., & Bédard, M. (2008). Older driver performance on a simulator: Associations between simulated tasks and cognition. *Advances in Transportation Studies*, Special Issue, 31–42.

Myers, R. S., Ball, K. K., Kalina, T. D., Roth, D. L., & Goode, K. T. (2000). Relation of useful field of view and other screening tests to on-road driving performance. *Perceptual and Motor Skills, 91*, 279–290.

Park, G. D., Cook, M. L., Rosenthal, T. J., Fiorentino, D., & Allen, R. W. (2006). Simulator assessment of older driver proficiency. *Advances in Transportation Studies*, Special Issue, 75–86.

Patomella, A. H., Tham, K., & Kottorp, A. (2006). P-drive: Assessment of driving performance after stroke. *Journal of Rehabilitation Medicine, 38*, 273–279.

Philip, P., Sagaspe, P., Taillard, J., Valtat, C., Moore, N., Akerstedt, T., et al. (2005). Fatigue, sleepiness, and performance in simulated versus real driving conditions. *Sleep, 28*, 1511–1516.

Reason, J. T., & Brand, J. J. (1975). *Motion sickness.* London: Academic Press.

Reed, M. P., & Green, P. A. (1999). Comparison of driving performance on-road and in a low-cost simulator using a concurrent telephone dialling task. *Ergonomics, 42*, 1015–1037.

Reimer, B., D'Ambrosio, L. A., Coughlin, J. E., Kafrissen, M. E., & Biederman, J. (2006). Using self-reported data to assess the validity of driving simulation data. *Behavior Research Methods, 38*, 314–324.

Riemersma, J. B. J., van der Horst, A. R. A., Hoekstra, W., Alink, G. M. M., & Otten, N. (1990). The validity of a driving simulator in evaluating speed-reducing measures. *Traffic Engineering and Control, 31*, 416–420.

Shinar, D., & Ronen, A. (2007). Validation of speed perception and production in a single screen simulator. *Advances in Transportation Studies*, Special Issue, 51–56.

Slick, R. F., Evans, D. F., Kim, E., & Steele, J. P. (2006, May). *Using simulators to train novice teen drivers: Assessing psychological fidelity as a precursor to transfer of training.* Paper presented at the Driving Simulation Conference—Asia/Pacific, Tsukuba, Japan.

Törnros, J. (1998). Driving behaviour in a real and a simulated road tunnel—A validation study. *Accident Analysis & Prevention, 30*, 497–503.

Toxopeus, R. (2007). *Differences between elderly and young drivers: Driving vigilance in two tasks.* Unpublished master's thesis, Lakehead University, Thunder Bay, Ontario, Canada.

Volkerts, E. R., van Laar, M. W., van Willigenburg, A. P. P., Plomp, T. A., & Maes, R. A. A. (1992). A comparative study of on-the-road and simulated driving performance after nocturnal treatment with lormetazepam 1 mg and oxazepam 50 mg. *Human Psychopharmacology: Clinical & Experimental, 7*, 297–309.

Wade, M. G., & Hammond, C. (1998). *Simulator validation: Evaluating driver performance in simulation and the real world* (Report No. 28). Minneapolis, MN: Minnesota Department of Transportation.

Wald, J., & Liu, L. (2001). Psychometric properties of the driVR: A virtual reality driving assessment. *Studies in Health Technology and Informatics, 81*, 564–566.

Watts, G. R., & Quimby, A. R. (1979). *Design and validation of a driving simulator for use in perceptual studies* (Report No. 907). Berkshire, UK: Transport and Road Research Laboratory.

Weaver, B., Bédard, M., McAuliffe, J., & Parkkari, M. (2009). Using the Attention Network Test to predict driving test scores. *Accident Analysis & Prevention, 41*, 76–83.

Yan, X., Abdel-Aty, M., Radwan, E., Wang, X., & Chilakapati, P. (2008). Validating a driving simulator using surrogate safety measures. *Accident Analysis & Prevention, 40*, 272–288.