



Editor's choice  
Scan to access more  
free content

# Simulated Driving Assessment (SDA) for teen drivers: results from a validation study

Catherine C McDonald,<sup>1,2,3</sup> Venk Kandadai,<sup>2</sup> Helen Loeb,<sup>2</sup> Thomas S Seacrist,<sup>2</sup> Yi-Ching Lee,<sup>2</sup> Zachary Winston,<sup>2</sup> Flaura K Winston<sup>2,3</sup>

<sup>1</sup>School of Nursing, University of Pennsylvania, Philadelphia, Pennsylvania, USA

<sup>2</sup>The Center for Injury Research and Prevention at The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA

<sup>3</sup>Department of Pediatrics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

## Correspondence to

Dr Catherine C McDonald, School of Nursing, University of Pennsylvania, Claire Fagin Hall, 418 Curie Boulevard, 414, Philadelphia, PA 19104-4217, USA; mcdonalc@nursing.upenn.edu

Received 3 November 2014

Revised 3 February 2015

Accepted 5 February 2015

Published Online First

4 March 2015

## ABSTRACT

**Background** Driver error and inadequate skill are common critical reasons for novice teen driver crashes, yet few validated, standardised assessments of teen driving skills exist. The purpose of this study is to evaluate the construct and criterion validity of a newly developed Simulated Driving Assessment (SDA) for novice teen drivers.

**Methods** The SDA's 35 min simulated drive incorporates 22 variations of the most common teen driver crash configurations. Driving performance was compared for 21 inexperienced teens (age 16–17 years, provisional license  $\leq 90$  days) and 17 experienced adults (age 25–50 years, license  $\geq 5$  years, drove  $\geq 100$  miles per week, no collisions or moving violations  $\leq 3$  years). SDA driving performance (Error Score) was based on driving safety measures derived from simulator and eye-tracking data. Negative driving outcomes included simulated collisions or run-off-the-road incidents. A professional driving evaluator/instructor (DEI Score) reviewed videos of SDA performance.

**Results** The SDA demonstrated construct validity: (1) teens had a higher Error Score than adults (30 vs 13,  $p=0.02$ ); (2) For each additional error committed, the RR of a participant's propensity for a simulated negative driving outcome increased by 8% (95% CI 1.05 to 1.10,  $p<0.01$ ). The SDA-demonstrated criterion validity: Error Score was correlated with DEI Score ( $r=-0.66$ ,  $p<0.001$ ).

**Conclusions** This study supports the concept of validated simulated driving tests like the SDA to assess novice driver skill in complex and hazardous driving scenarios. The SDA, as a standard protocol to evaluate teen driver performance, has the potential to facilitate screening and assessment of teen driving readiness and could be used to guide targeted skill training.

## INTRODUCTION

In the USA in 2011, almost 2000 drivers between the ages of 15 years and 20 years died in MVCs, and an additional 180 000 young drivers suffered injuries, making MVCs the leading health threat to adolescents.<sup>1</sup> Poor driving skill, such as failure to recognise hazards, inattention to the forward roadway, and driving too fast for local conditions, contributes to nearly three-quarters of teen driver crashes.<sup>2–3</sup> Even though we can identify major contributors, methods to reduce crashes among novice teen drivers have been hampered by a lack of valid assessments that measure driving skill, particularly in complex driving situations. Given that teen drivers account for a disproportionately high percentage of crashes,<sup>4</sup> and the resultant morbidity

and mortality exerting a societal burden,<sup>5</sup> validated, standardised assessment of teen driving skill is needed.

On-the-road assessments of driving performance with evaluators,<sup>6</sup> while important and necessary, are limited in hazard exposure, and include threats to validity. Key factors, such as traffic and environmental conditions, cannot be replicated, and study designs cannot place teens in dangerous on-the-road driving situations. However, teens crash in complex and dangerous scenarios;<sup>7–8</sup> therefore, assessment of how teens perform in these situations can provide important information for prevention strategies. Driving simulators could offer a replicable, safe complement to on-the-road evaluations;<sup>9</sup> however, no validated driving simulator protocols to assess teen driver performance exist. Our goal was to create a standardised, replicable simulated assessment of teen driving performance in scenarios known to be associated with teen crashes, delivered within a driving simulator's safe environment.

We have undertaken a series of systematic steps to develop an assessment of teen driver performance: the Simulated Driving Assessment (SDA). The scenarios and measurement strategies in the SDA were theoretically grounded in actual crash data. The SDA was designed to expose drivers to complex driving scenarios that are often the cause of serious crashes and to measure their driving performance via key metrics determined from the empirical literature and expert opinion.<sup>10–17</sup> Crash data from the National MVC Causation Survey (NMVCCS) were used to determine the SDA-simulated scenarios.<sup>18–19</sup> The top three types of serious crashes (rear-end collisions, left turn intersection collisions, and right side run-off-the-road events) represented approximately 30% of NMVCCS teen crash configurations.<sup>19</sup> Variations of these three types of serious crashes became the driving scenarios in the SDA so that the assessment would approximate the common real-world scenarios in which teen drivers crash; that is, have ecological validity.<sup>19</sup> Initial analyses were carried out to help establish the usefulness of the SDA. We examined the potential for learning effects during the SDA and found no training effect or induced hypervigilance because of exposure to multiple potential crash scenarios in the SDA.<sup>20</sup> Given that a driving evaluator instructor (DEI) often serves as the best available gold standard for assessing driving performance, we compared DEI-rated drivers' performance during the SDA (a derived categorisation of skill/less skilled) to history of police-reported crashes. Seven of the eight (87.5%) experienced adult drivers who had at



CrossMark

**To cite:** McDonald CC, Kandadai V, Loeb H, et al. *Inj Prev* 2015;**21**:145–152.

least one police-reported crash were categorised as less skilled by the DEI, resulting in a sensitivity of 0.87 (95% CI 0.47 to 0.99).<sup>21</sup>

Given this foundation, the purpose of this study was to assess the construct and criterion validity of the SDA for novice teen drivers, crucial tests of its usefulness for driver assessment. In order to evaluate the construct validity of the SDA, we: (1) compared safe driving performance measures among *teens* and *adults* (Error Score); and (2) assessed the ability of the SDA to predict Simulated Negative Driving Outcomes from the Error Score. In order to evaluate the SDA's criterion validity, we compared the Error Score to an expert criterion, an independent video review of SDA performance by a professional driving evaluator/instructor (DEI Score).

## METHODS

### SDA procedures

The methods to develop the SDA scenarios have been described above. Prior to completing the SDA, participants drove an unscored familiarisation segment (~7–10 min) to adjust to simulator dynamics. The SDA lasted approximately 35–40 min and included three modules that were randomised (by a random number generator). During the SDA, participants were exposed to 22 variations of the most common teen driver crash configurations (rear-end collisions, left turn intersection collisions, and right side run-off-the-road events).<sup>19</sup> The 22 crash scenarios were distributed across the three modules, separated by intervening straight roads, curves and turns not intended to trigger collisions. The length between scenarios ranged between about 300 m and 1800 m, with an average of about 930 m. If the participant drove the SDA safely, crashes were avoidable. In order to decrease the potential for a learning effect, no feedback on performance was given to participants during or after the drive.<sup>20</sup>

### SDA driving performance metrics

The 14 driving performance metrics scored during the SDA (determined by the empirical literature and expert opinion) represent key measurable driving behaviours, including basic (eg, use of turn signals) and advanced (eg, presence of hazard anticipation glances) manoeuvres. Table 1 outlines the definitions, calculations, sources of data and criteria for errors in each of the metrics, all of which were determined by on-the-road and/or simulator studies. Table 2 outlines the potential crash scenarios, number of occurrences in the SDA and errors (dichotomous) scored in the 22 potential crash scenarios.

### Apparatus

The Realtime Technology, Inc. (RTI) fixed-based driving simulator system used in this study included a driver seat, three-channel 46" LCD (liquid-crystal display) panels (160° field of view), rear view, left and right mirror inlayed images, active pedals and a steering system. Visual rendering and graphics were delivered at 1280×1024 resolution at 60 Hz. Raw simulator data (eg, velocity, position) were collected at 60 Hz and were reduced to the 14 driving performance metrics as outlined in table 1. Videos of the participant's driving behaviour in the simulator were also recorded by three cameras (positioned over the right shoulder, pointing at the participant's face and over the foot for view of brake and accelerator). Video playback data were assessed by the DEI (further described below).

Applied Science Laboratories' (ASL) Mobile Eye was used to capture eye movements on the forward scene at 30 Hz. Participants wore a pair of glasses (or alternative goggles worn

over eyeglasses) with two video cameras (forward image and right eye of participant) to capture gaze location during the SDA. Video data from the two cameras were integrated into a single video with superimposed cross-hairs for eye gaze location. Videos were used for coding in analysis as outlined in table 1.

### Participants

Two groups were enrolled: (1) inexperienced teens (age 16–17 years, Pennsylvania (PA) provisional license ≤90 days); and (2) experienced adults (age 25–50 years, PA license ≥5 years, drove ≥100 miles per week, no collisions or moving violations ≤3 years).

Individuals were excluded for a self-report history of migraines or motion sickness, current pregnancy, or non-English speaking. Teen participants were recruited via mailings from paediatric primary care facilities, driving schools and word of mouth. Experienced adults were recruited via flyers, the Wharton Behavioral Lab at the University of Pennsylvania and word of mouth. Participants were compensated \$50. Study procedures were approved by the institutional review boards at The Children's Hospital of Philadelphia and an administrative agreement with the University of Pennsylvania.

### Main measures

#### Demographic data

Self-report data on age, gender, race/ethnicity and length of licensure were collected.

#### Error Score

Error Score was based on errors in 14 driving performance metrics in the 22 potential crash scenarios. Only errors relevant to a specific scenario were calculated (eg, turning from a correct lane was not scored in the rear-end events; see table 1 for calculations and table 2 for scoring). A minimum of 1 and maximum of 10 driving performance metrics were scored in each of the 22 scenarios. A sum score of errors in driving performance metrics was aggregated across potential crash scenarios (Error Score). Potential range of the Error Score was 0–117.

#### Simulated negative driving outcome (simulated collisions and run-off-the-road)

Each of the 22 scenarios had potential for a negative driving outcome, either a simulated collision or run-off-the-road incident. Only one simulated negative driving outcome could be assigned per scenario, whichever occurred first. Simulated collisions were defined as an overlap of the participant's vehicle with other vehicles programmed for the potential crash scenario. Collisions were derived from simulator data on position, orientation and dimensions of the participant and nearest vehicle, and were verified by video review. For five participants, automated simulator data for collisions were unavailable due to a programming error and were independently video coded and reconciled for 100% agreement by two research team members. Run-off-the-road incidents were defined as any instance of the vehicle perimeter departing the roadway (either side). Run-off-the-road incidents were derived from simulator data of participant vehicle position, orientation and dimensions. Possible range of total SDA simulated negative driving outcomes was 0–22.

#### DEI Score

A professional DEI with 20 years of driver training experience reviewed a four-quadrant divided screen of SDA video playback of the three scored modules (familiarisation segment excluded).

**Table 1** Driving performance metric definitions, calculations and criterion for error

Driving performance metric	Definition of metric	Calculation	Source	Criterion for error in SDA
<b>Operational manoeuvres</b>				
Lane choice <sup>10</sup>	Lane selection during turn at an intersection	Lane position up to 2 s prior to initiation of turn as indicated by a turn of the steering wheel beyond $\pm 45^\circ$	Simulator	Controlled and uncontrolled left turns: turning left at intersection from the right lane
Lateral acceleration during turn <sup>11</sup>	Lateral acceleration during a turn at an intersection	Lateral acceleration	Simulator	Controlled and uncontrolled left turns: lateral acceleration $\geq 0.5$ g
Speed <sup>10</sup>	Velocity	Velocity relative to posted speed limit	Simulator	Controlled left turns, rear-end and hidden hazard (car in driveway): velocity exceeds posted speed limit by 10% in clear conditions and 5% in rainy conditions. Curves: velocity exceeds posted speed limit at any point in the curve. Uncontrolled left turns: velocity exceeds posted speed limit at any point in the turn
Stopping behaviour at a stop sign <sup>10</sup>	Velocity at a stop sign	Minimum velocity at stop sign	Simulator	Controlled left turns: failure to reach velocity of zero at a stop sign
Traffic check <sup>10</sup>	Glance sequence at intersection	Video coder identified glance location sequence	Eye tracking data-video	Controlled left turns: failure to make a glance sequence of left-right-left within 5.5 m of intersection
Turn signal use <sup>10</sup>	Application of turn signal to alert other road users of a turn	Activation of turn signal	Simulator	Controlled and uncontrolled left turns: failure to use turn signal
Yaw <sup>11</sup>	Vehicle heading correction after completion of a turn at an intersection	Vehicle heading correction after completion of turn as indicated by the return of the steering wheel to $0^\circ$	Simulator	Controlled and uncontrolled left turns: vehicle heading correction is altered by more than $\pm 6^\circ$ after turn completion
<b>Advanced manoeuvres</b>				
Braking behaviour in a hazardous situation <sup>11–13 16</sup>	Instance of substantial longitudinal deceleration in response to a hazardous situation	Rear-end events: based on headway time and instance of longitudinal deceleration $>0.6$ g prior to a collision. Hidden hazard (car in driveway): based on velocity and instance of longitudinal deceleration $>0.6$ g prior to a collision	Simulator	Rear-end: headway time $<3$ s and no instance of longitudinal deceleration $>0.6$ g Hidden hazard (car in driveway): Velocity exceeds posted speed limit by 5% and no instance of longitudinal deceleration $>0.6$ g
Excess glance duration $>1.5$ s <sup>14</sup>	Summed glance duration off forward roadway over 1.5 s	Time duration between when the cross-hairs of the ASL eye-tracker left the area of the centre display view and reappeared during a distraction task: duration of glances were summed	Eye tracking data video	Rear-end with distraction: glance duration off forward roadway $>1.5$ s
Gap selection <sup>17</sup>	Choice of time to enter an intersection and proximity to other vehicle in intersection	Video coding determined if participant waited for cross traffic. For participants who did not wait, postencroachment time (PET) was calculated	Simulator (video and data)	Controlled left turns: PET $<1.5$ s. Participants who waited for cross traffic did not receive an error
Headway time <sup>16</sup>	Instantaneous time to reach lead vehicle	Distance/velocity of participant vehicle centre of gravity to lead vehicle centre of gravity. Video coder estimated distance for a minority of participant scenarios where vehicle information was not available*	Simulator	Rear-end: distance/velocity of participant vehicle to lead vehicle $<3$ s
Scan left for intersection <sup>15</sup>	Glance to the left as participant approaches an intersection	Video coder identified glances to left as participant approached intersection: coded from the location of instructions to make a left turn until $\sim 5.5$ m from intersection	Eye tracking data video	Controlled and uncontrolled left turns: no recorded glances to left as participant approached intersection
Scan right for intersection <sup>15</sup>	Glance to the right as participant approaches an intersection	Video coder identified glances to right as participant approached intersection: coded from the location of instructions to make a left turn until $\sim 5.5$ m from intersection	Eye tracking data video	Controlled and uncontrolled left turns: no recorded glances to right as participant approached intersection
Scan right for hidden hazard <sup>15</sup>	Glance to the right as participant approaches a hidden hazard	Video coder identified glances on fence, driveway and car as participant approached hazard event and record time stamps of glances and when car pulls out	Eye tracking data video	Hidden hazard (car in driveway): no recorded glances to right as participant approached hidden hazard event

\*Video coding calculation of headway time to known simulator derived values of headway time and found that the mean difference between both samples was 0.13 s (95% CI 0.01 to 0.27).

SDA, Simulated Driving Assessment.

The video included forward roadway, right foot behaviour (brake and throttle), and over the participant's right shoulder and face. Speed in miles per hour (mph) and left and right turn signals were displayed; however, no simulator or eye-tracking data or selected measures for scoring were provided to the DEI. The DEI was blinded to driver age and SDA Error Score. The

DEI rated driving skill (scale 0–10; 0=unskilled and 10=proficient) in eight domains: attention maintenance, communication and right of way, gap selection, hazard anticipation and response, managing blind spot, road positioning, speed management and vehicle control. As previously stated, this DEI demonstrated a high sensitivity in identifying skill level in the SDA as

**Table 2** Scoring of driving performance metrics in potential crash scenario type

Potential crash scenario type	Variations	Number in SDA	Driving performance metric												
			Lane choice	Lateral acceleration during a turn	Speed	Stopping behaviour at a stop sign	Traffic check	Turn signal	Yaw	Braking behaviour in a hazardous situation	Excess glance duration	Gap selection	Headway time	Scan-left for intersection	Scan-right for intersection
Controlled left turns	Three-way intersection with cross-traffic, no obstruction	2	X	X	X	X	X	X	X	X		X	X	X	
	Three-way intersection with cross-traffic, with an obstruction	2	X	X	X	X	X	X	X		X	X	X	X	
	Four-way intersection with cross-traffic, no obstruction	2	X	X	X	X	X	X	X		X	X	X	X	
	Out of a parking lot	2		X	X		X	X	X		X				
Curves	Curves in rainy weather	3			X										
Rear-end—no distraction	Lead vehicle stops suddenly in the left lane in order to turn left	2			X					X			X		
	Lead truck in right lane stops suddenly	3			X					X			X		
Rear-end— with distraction	Lead truck in right lane stops suddenly while participant looks for loose change or tunes the radio	2			X					X	X*		X		
Hidden hazard (car in driveway)	Car emerges from the right from an obstructed driveway in rainy weather	2			X					X					X
Uncontrolled left turns	Three-way intersection where participant has a green light	2	X	X	X					X	X		X	X	

\*Excess glance duration was only scored in 1 distraction event.  
SDA, Simulated Driving Assessment.

related to previous crash history.<sup>21</sup> A sum score was created (DEI Score) with a possible range of 0–80.

## Analysis

Table 1 outlines metric definitions, calculations, sources of data and criteria for errors. Custom MATLAB (Mathworks, Natick, Massachusetts, USA) code was used to reduce raw simulator data. For categorical variables, frequencies and percentages were computed. For continuous variables, median, IQR, and full range were computed. Construct validity was examined by assessing the following hypotheses: (H<sub>1</sub>) teen drivers would have a higher Error Score than adults; and (H<sub>2</sub>) higher Error Score would predict increased propensity for a simulated negative driving outcome. The Wilcoxon rank-sum test and Fisher's exact test were used to compare teen and adult performance. A modified Poisson regression was used to estimate the RR and 95% CI to predict the total sum of simulated negative driving outcomes based on experience (inexperienced teen or experienced adult) and Error Score. Criterion validity was examined by assessing the following hypothesis: (H<sub>3</sub>) DEI score would be negatively associated with Error Score (eg, participants rated as more proficient by the DEI would have fewer errors or a lower Error Score). Spearman correlation was used to determine the association between Error Score and DEI Score. The videos viewed by the DEI included obvious collisions and run-off-the-road instances; therefore, the association between DEI score and negative driving outcomes was not examined. For hypothesis testing, the Type-I error rate was set at  $\alpha=0.05$ . All statistical analyses were performed using R v3.0.1 (<http://www.r-project.org>).

## RESULTS

Our analytic sample included 38 participants who completed the SDA (see figure 1). The teen group included 21 participants: 33% female, 95.2% Caucasian, 4.8% African-American, 100% non-Hispanic, median age=16 years (IQR 16–17; range 16–17), with provisional license for a median of 35 days (IQR 7–44; range 1–88). The adult group included 17 participants: 41% female, 70.6% Caucasian, 23.5% African-American, 5.9% race not reported, 100% non-Hispanic, median age=34 years (IQR 28–42; range 25–50), licensed for a median of 15 years (IQR 11–22; range 8–33). There were no statistically significant differences between the teen and adult groups across sex, race

and ethnicity. As expected, statistically significant differences between the teens and adults were observed for age ( $Z=5.40$ ,  $p<0.001$ ) and length of licensure ( $Z=5.23$ ,  $p<0.001$ ) (Wilcoxon rank-sum tests).

Table 3 includes the SDA Error Score and errors in each driving performance metric for teens, adults and the total sample. For the total sample, the Error Score range was 4–54 errors (teen range, 13–54; adult range, 4–50). Consistent with H<sub>1</sub>, teens had a higher Error Score than did the adults ( $Z=-2.37$ ,  $p=0.02$  Wilcoxon rank-sum test). Note, median and IQRs of errors and frequency of teens and adults with  $\geq 1$  error in each driving performance metric during the entire SDA are also reported in table 3; however, given the small sample, we consider comparisons between teens and adults exploratory and, thus, do not report p values.

During the SDA, 42.9% of teens and 29.4% of adults had at least one simulated negative driving outcome. Total sample range was 0–4 (teen range 0–3 and adult range 0–4); total sample median was 0 (IQR 0–2) (teens median 0 (IQR 0–2); adult median 0 (IQR 0–1)). There were no statistically significant differences in total simulated negative driving outcomes ( $Z=-0.80$ ,  $p=0.43$  Wilcoxon rank-sum test) between teens and adults or percentage of teens and adults with at least one simulated negative driving outcome ( $p=0.51$ , Fisher's exact test). Note: among negative driving outcomes, 94% were simulated collisions.

Consistent with H<sub>2</sub>, across all participants, for every one additional driving error, there was an 8% increase in the risk of a simulated negative driving outcome (RR=1.08; 95% CI 1.05 to 1.10,  $p<0.01$ ). Figure 2 plots the exponential relationship between the propensity to have a simulated negative driving outcome (mean of lambda) and Error Score. Using a Wald test of the coefficients ( $Z=-0.36$ ,  $p=0.72$ ) and Akaike Information Criterion (AIC), experience was rejected as a covariate. Experience alone was not a significant predictor of simulated negative driving outcome (RR=1.38; 95% CI 0.67 to 3.04).

DEI score ranged from 6 to 66 (teen range 23–66; adult range, 6–65) with a median of 46 and IQR 32–59 (teen median=46, IQR 34–58; adult median=46; IQR 32–61). There was no statistically significant relationship between DEI score and experience ( $Z=-0.25$ ,  $p=0.80$ ). However, consistent with H<sub>3</sub>, Error Score was significantly associated with DEI Score ( $r=-0.66$ ,  $p<0.001$ ).

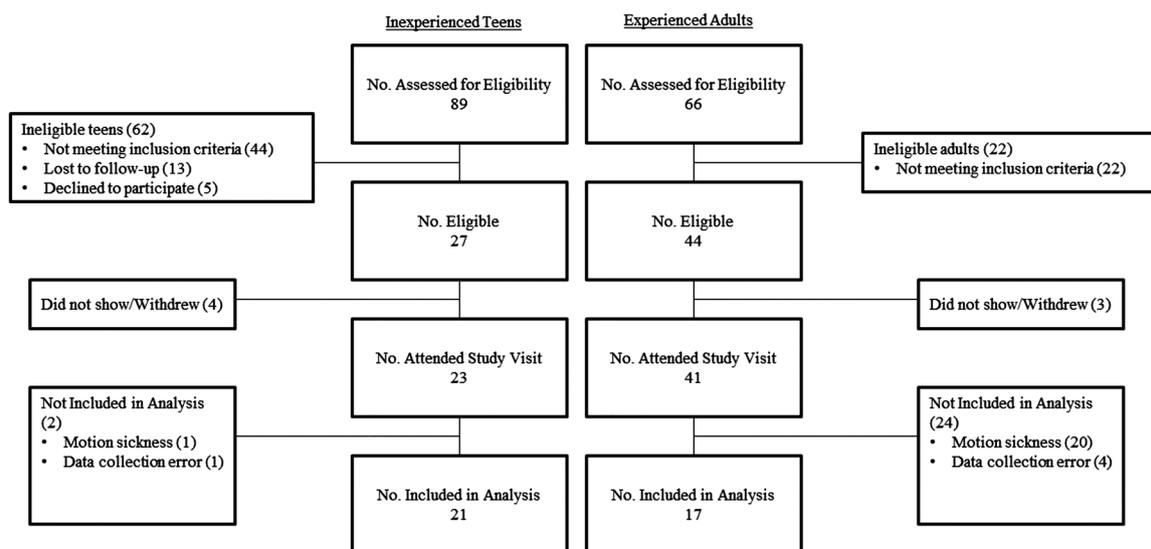


Figure 1 Derivation of analytical sample.

**Table 3** Error Score, errors in each driving performance metric and simulated negative driving outcomes

	Possible range	Total sample (n=38) Median (IQR)	Teens (n=21)		Adults (n=17)		p Value
			Median (IQR)	Teens with $\geq 1$ error in each driving performance metric n (%)	Median (IQR)	Adults with $\geq 1$ error in each driving performance metric n (%)	
Error score	0–117	22.5 (13–36)	30 (19–36)		13 (10–32)		0.02*
Driving performance metrics							
Lane choice	0–8	0 (0–0)	0 (0–0)	5 (23.8)	0 (0–0)	1 (5.9)	
Lateral acceleration during a turn	0–10	2 (0–6)	2 (1–6)	16 (76.2)	2 (0–5)	11 (64.7)	
Speeding	0–22	4 (2–7)	6 (2–7)	19 (90.5)	3 (0–4)	12 (70.6)	
Stopping behaviour at a stop sign	0–6	0 (0–1)	0 (0–3)	8 (38.1)	0 (0–1)	7 (41.2)	
Traffic check	0–8	3 (1–4)	3 (2–4)	14 (66.7)	2 (0–3)	16 (94.1)	
Turn signal	0–10	0 (0–0)	0 (0–0)	0 (0)	0 (0–0)	3 (17.6)	
Yaw	0–10	1 (0–1)	0 (0–1)	11 (52.4)	1 (0–2)	12 (70.6)	
Braking behaviour in a hazardous situation	0–9	2 (1–3)	3 (2–4)	21 (100)	1 (0–2)	12 (70.6)	
Excess glance duration	0–1	0 (0–0)	0 (0–1)	6 (28.6)	0 (0–0)	3 (17.6)	
Gap selection	0–8	6 (4–8)	6 (5–8)	8 (38.1)	6 (4–7)	8 (47.1)	
Headway time	0–7	3 (2–5)	4 (3–5)	21 (100)	2 (1–3)	13 (76.5)	
Scan left for intersection	0–8	3 (1–4)	4 (3–5)	20 (95.2)	1 (0–3)	15 (88.2)	
Scan right for intersection	0–8	3 (2–5)	4 (2–5)	21 (100)	3 (2–4)	11 (64.7)	
Scan hidden hazard	0–2	0 (0–0)	0 (0–1)	7 (33.3)	0 (0–0)	2 (11.8)	
Simulated negative driving outcomes	0–22	0 (0–2)	0 (0–2)		0 (0–1)		0.43

Note: Error Score is the sum score of errors in driving performance metrics aggregated across potential crash scenarios. Teens and adults had similar distributions of individual errors, error score and simulated negative driving outcomes. Missing Data: For a minority of cases with device calibration failure, eye-tracker-derived measures could not be calculated for the hazard anticipation glances, traffic check and summed excess glance duration (missing data for each measure by scenario ranged from 0% to 21%). Missing data for hazard anticipation glance or traffic check data were imputed from available data in the other scenarios. Given that there was only one event scored for summed excess glance error, missing data for this driving performance error were assigned a non-error.

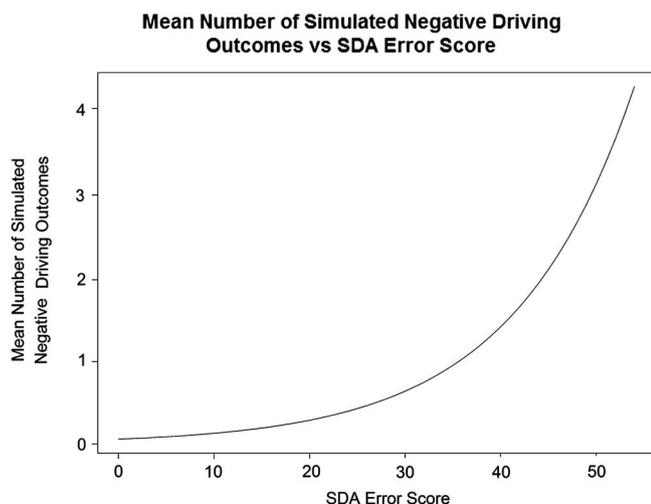
## DISCUSSION

The results provide support for the construct and criterion validity of the SDA in its ability to differentiate driver performance by use of established safety metrics and to predict simulated negative driving outcomes. Teen drivers within 3 months of licensure made more total errors in the SDA than did experienced adult drivers. Furthermore, the SDA demonstrated a range in driver performance errors and a near linear relationship between Error Score and propensity to have at least one simulated negative driving outcome; with each error in the SDA, the risk for crashing or running off the road increased 8%

(RR=1.08; 95% CI 1.05 to 1.10,  $p < 0.01$ ). Evidence for criterion validity is provided by the association between independent expert rating of driving performance and calculated scores of safety measures derived from simulator and eye-tracker data ( $r = -0.66$ ,  $p < 0.001$ ).

The SDA Error Score, but not ratings from expert video review, showed a difference in driving performance between teen and adult drivers. The lack of association of video review and experience was likely due to potential limitations in video review for assessment of performance. Even though a DEI is often considered the best available standard for evaluation, the position of the DEI in a vehicle may limit ability to fully assess skill. For example, from the video review, the DEI was not able to precisely see where the participants' were looking (ie, eye glance location) which is similar to the situation with on-road assessments where an evaluator's view may be restricted. Given that poor scanning contributes to novice driver crashes,<sup>2 3</sup> a simulator and eye tracker can play an important role in assessing crucial skills like hazard anticipation and glance duration off the forward roadway.

A key goal of the SDA is to differentiate between skilled and non-skilled drivers; thus, we used experience as a proxy for skill. Although adults in our sample were experienced, they did not necessarily perform like skilled drivers. This is consistent with our previous research which revealed that the experienced adults, while eligible for the study, may have included drivers who were less skilled and less safe drivers: 8 of the 17 adults had at least one police-reported crash during their PA licensure, none that occurred during the 3-year window used in the study exclusion criteria.<sup>21</sup> (Note: none of the inexperienced group had a police-reported crash prior to enrollment in the study). Therefore, our adult driver group contained some drivers who did not demonstrate the 'gold standard' for safe driving. This



**Figure 2** Mean number of simulated negative driving outcomes vs SDA Error Score.

may, in part, explain why a statistical difference between the adult and teen group was found in Error Score, but we did not find that the SDA differentiated simulated negative driving outcomes based on experience. There was a trend, however, to suggest that teens may have more simulated negative driving outcomes than adults (42.9% vs 29.4%; RR=1.38, 95% CI 0.67 to 3.04,  $p>0.05$ ). With the narrow distribution of the continuous count of crashes and dichotomous outcome variable (no crashes vs  $\geq 1$  crashes), analysis of crash counts was likely underpowered to detect statistical differences. Future validation of the SDA (and other studies that use adults as a skilled/safe driver comparison group) may need to use a larger sample and additional criteria that enhance exclusion of unskilled, unsafe drivers.

Some of our individual driving performance metrics had a median of 0 errors by teens or adults (eg, turn signal, lane choice). These metrics were included to provide face validity of the new SDA tool as these are common metrics used by driving assessors as part of licensing. However, future iterations of the SDA scoring may choose to refine or eliminate these metrics due to lack of variation among participants. Other metrics proved more variables, and demonstrated the usefulness of the tool and also the range of driving skill among the participants. Additionally, future research with a larger sample should explore how each error metric differs among teens and adults.

By examining multiple performance measures in scenarios designed to replicate the most common crash configurations, the SDA provides information about skill deficits in hazardous situations. These findings demonstrate the SDA's usefulness as a measure of driving skills known to be involved in teen driver crash risk. Teens displayed mastery of basic skills, such as using turn signals, but more advanced driving behaviours proved challenging. Teens had repeated errors of advanced behaviours, including braking in hazardous situations and hazard anticipation glances, and demonstrated repeated negative driving outcomes during the SDA. These results complement those from an evaluation of an on-road assessment demonstrating teens' ability to perform basic manoeuvres but had deficits in more advanced, critical driving behaviours.<sup>22</sup>

This study is not without limitations. The simulated driving environment replicates but is not identical to on-the-road driving; however, the SDA's advantage is to safely expose teens to high-risk driving situations. Future prospective studies with larger samples should examine the ability of the SDA to predict teen driver on-road behaviour and crashes. The study inclusion/exclusion criteria limit the applicability of the results to the populations that were studied. Larger samples are needed to generalise results to the broader population or to special populations with medical and psychiatric conditions. The metrics for scanning (eg, scan left for intersection) were not an exact replication of scenarios and measurement of the hazard anticipation glances in the Fisher *et al*<sup>15</sup> work. However, the emphasis of this metric on scanning in potentially hazardous situations was the basis for selection. Future work should further examine measurement strategies for scanning in the SDA. Incomplete performance data were available for 15 adults and 1 teen who stopped the SDA due to simulator sickness (see figure 1 for determination of sample). Overall, the percentage of adults with simulator sickness in our study (36.6%) is consistent with reports in the literature.<sup>12</sup> For the stated purpose of the SDA—assessment of novice teen driver performance—the SDA resulted in minimal simulator sickness. This study focused on assessment of driving skill in relationship to experience as a marker of safe driving. A more comprehensive assessment of

driving risk should also include risky driving and behavioural tendencies. A single-rater (DEI) with over 20 years of experience in driver training was used to assess video recordings, which was only one form of validity. In the future, the use of multiple raters might be of value, although we confirmed the ability of this DEI to accurately identify unskilled, unsafe drivers.<sup>21</sup>

## CONCLUSIONS

This study supports the validity of the SDA for use in a safe and replicable manner to assess novice driver skill in complex and hazardous driving scenarios. The SDA, as a standard protocol to assess teen driver performance, could facilitate screening and assessment of teen driving readiness, and could be used to guide targeted skill training.

### What is already known on the subject?

- ▶ Driver error and inadequate skill are the main contributors to teen driver crashes.
- ▶ On-road assessments of driving performance are limited due to limited exposure (safety concerns) and replicability (variable traffic/environmental conditions).
- ▶ There is a gap in validated assessments of teen driver skill, particularly in complex driving situations that have inherent risk.
- ▶ Advanced driving simulators provide a realistic yet safe complement to on-road tests.

### What this study adds?

- ▶ This study demonstrated validity of the Simulated Driving Assessment (SDA) for safe evaluation of skill in high-risk driving scenarios.
- ▶ A tool like the SDA could inform teen driving evaluations and guide individualised driver training and management plans.

**Acknowledgements** We would like to acknowledge the Simulator Program at the Center for Injury Research and Prevention, directed by Yi-Ching Lee, PhD, and coordinated by Dana Bonfiglio; Jason Tanenbaum, Dana Bonfiglio, and student research assistants for their role in data collection and analyses; Donald Fisher, PhD, and Daniel Mayhew, MA for their assistance in determining the domains and scenarios and Matthew Romoser, PhD, for his assistance in programming the scenarios; Michael Elliot, PhD, for his statistical guidance; Marilyn S Sommers, PhD, RN and Nancy Kassam-Adams, PhD for their guidance during the study and thoughtful review of the manuscript; Gerald Murphy for his assistance with this study; the Wharton Behavior Lab; Pediatric Research Consortium at The Children's Hospital of Philadelphia; and the study participants without whom the research would not have been possible.

**Contributors** CCM, Y-CL and FKW contributed to the conception and design of the study, acquisition of data, and analysis and interpretation of data; drafting and critical review of the manuscript; and approval of the final manuscript. Kandadai contributed to the acquisition of data, and analysis and interpretation of data; drafting and critical review of the manuscript; and approval of the final manuscript. HL, TSS and ZW contributed to the analysis and interpretation of data; drafting and critical review of the manuscript; and approval of the final manuscript.

**Funding** This project is funded, in part, under a grant with the Pennsylvania Department of Health (PI: Flaura K. Winston). The department specifically disclaims responsibility for any analyses, interpretations or conclusions. Dr Catherine C McDonald was supported by the National Institute of Nursing Research of the National Institutes of Health under Award Number K99NR013548. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors would also like to acknowledge the National Science Foundation (NSF) Center for Child Injury Prevention Studies at the Children's Hospital of Philadelphia (CHOP) for partial

support of Flaura K Winston (NSF Grant ID: 1034593) and for student support (NSF Grant ID: 1451897) for this study. The views presented are those of the authors and not necessarily the views of CHOP or the NSF.

**Competing interests** None.

**Ethics approval** The Children's Hospital of Philadelphia and an administrative agreement with the University of Pennsylvania.

**Provenance and peer review** Not commissioned; externally peer reviewed.

## REFERENCES

- National Highway Traffic Safety Administration. *Traffic safety facts 2011 data: young drivers*. Washington DC 2013 [16 January 2015]. <http://www-nrd.nhtsa.dot.gov/Pubs/811744.pdf>
- Braitman KA, Kirley BB, McCart AT, et al. Crashes of novice teenage drivers: characteristics and contributing factors. *J Safety Res* 2008;39:47–54.
- Curry AE, Hafetz J, Kallan MJ, et al. Prevalence of teen driver errors leading to serious motor vehicle crashes. *Accid Anal Prev* 2011;43:1285–90.
- Insurance Institute for Highway Safety. *Fatality facts 2010: teenagers*. Arlington, VA: Insurance Institute for Highway Safety, Highway Loss Data Institute; 2010 [16 January 2015]. <http://www.iihs.org/iihs/topics/t/teenagers/fatalityfacts/teenagers>
- Centers for Disease Control and Prevention. Teen driver fact sheet. 2013 [16 January 2015]. [http://www.cdc.gov/motorvehiclesafety/teen\\_drivers/index.html](http://www.cdc.gov/motorvehiclesafety/teen_drivers/index.html)
- Romanowicz PA, Hagge RA. *An evaluation of the validity of California's driving performance evaluation road test*. Sacramento, CA. 1995 [16 January 2015]. [http://apps.dmv.ca.gov/about/profile/rd/r\\_d\\_report/Section\\_2/S2-154.pdf](http://apps.dmv.ca.gov/about/profile/rd/r_d_report/Section_2/S2-154.pdf).
- Bingham CR, Ehsani JP. The relative odds of involvement in seven crash configurations by driver age and sex. *J Adolesc Health* 2012;51:484–90.
- McDonald CC, Curry AE, Kandadai V, et al. Comparison of teen and adult driver crash scenarios in a nationally representative sample of serious crashes. *Accid Anal Prev* 2014;72:302–8.
- Boyle LN, Lee JD. Using driving simulators to assess driving safety. *Accid Anal Prev* 2010;42:785–7.
- Pennsylvania Department of Transportation. Pennsylvania's Driver's Manual: Chapter 3. [16 January 2015]. [http://www.dmv.state.pa.us/pdotforms/pa\\_driversmanual/chapter\\_3.pdf](http://www.dmv.state.pa.us/pdotforms/pa_driversmanual/chapter_3.pdf)
- Simons-Morton BG, Ouimet MC, Zhang Z, et al. Crash and risky driving involvement among novice adolescent drivers and their parents. *Am J Public Health* 2011;101:2362–7.
- Fisher DL, Rizzo M, Caird J, Lee JD. eds. *Handbook of driving simulation for engineering, medicine, and psychology*. Boca Raton, FL: CRC Press, Taylor & Francis Group; 2011.
- Romoser M, Deschamps M, Wilson H, et al. Investigating differences between experienced adult drivers and teen drivers with low-cost vehicle data recorder. *Transp Res Rec* 2012;2321:79–87.
- Pollatsek A, Divekar G, Fisher D. New measure of inattentiveness to forward roadway. *J Transportation Res Board* 2012;2321:31–7.
- Fisher DL, Pollatsek AP, Pradhan A. Can novice drivers be trained to scan for information that will reduce their likelihood of a crash? *Inj Prev* 2006;12(Suppl 1): i25–9.
- New South Wales Government. *A guide to the driving test*. Transport Road and Maritime Services, 2010.
- Yasar A-U-H, Berbers Y, Preuveneers D. A computational analysis of driving variations on distributed multiuser driving simulators. *Proceedings of the 19th IASTED International Conference on Modelling and Simulation*, 2008:178–86.
- National Highway Traffic Safety Administration. *National motor vehicle crash causation survey: report to Congress*. Washington DC: National Highway Traffic Safety Administration; 2008 [16 January 2015]; <http://www-nrd.nhtsa.dot.gov/Pubs/811059.PDF>
- McDonald CC, Tanenbaum JB, Lee Y-C, et al. Using crash data to develop simulator scenarios for assessing young novice driver performance. *Transp Res Rec* 2012;2321:73–8.
- McDonald CC, Lee YC, Tanenbaum JB, et al. Evaluation of unintended training effects from collision scenarios in a simulated assessment for young novice drivers. *Transportation Research Board 92nd Annual Meeting Compendium of Papers, Paper No 13-2794*; Washington, DC, 2013.
- Winston FK, McDonald CC, Kandadai V, et al. Experience and skill predict failure to brake errors: further validation of the simulated driving assessment. *SAE Technical Paper 2014-01-0445*. 2014.
- Mirman J, Lee Y-C, Kay J, et al. Development of web-based parent support program to improve quantity, quality, and diversity of teens' home-based practice driving. *Transp Res Rec* 2012;2318:107–15.

### Is injury proneness genetic?

A review paper in *Clin J Sports Medicine* suggests there may be a genetic basis for repeated injuries. One focus is on genes that control collagen production. Athletes who tore their anterior cruciate ligament were four times as likely as uninjured controls to have a blood relative who had the same injury. Another possible genetic marker relates to bone mineral density. Some evidence suggests that osteoporosis in older women and increased stress fractures in young women may be familial. (Noted by IBP)

### Project Zero: towards injury-free India

In Mangalore, India, a city with a population of nearly one half million, the University has launched Project Zero, with the goal of eliminating all injury deaths in an unspecified period. The strategy is to "appeal to every responsible citizen of India to join and pledge road safety in mind and body, heart and soul". It notes that "present efforts ... to enforce law(s) and to deliver post crash treatment is minimal". The rhetoric urges "Safe and slow driving, road manners, wearing helmet, seat belt and other necessary precautions ..." (Noted by IBP)

### New Ford includes many crash prevention features

The new hi-tech Ford S-Max car includes >20 injury prevention technologies as well as an Intelligent Speed Limiter. (Noted by IBP)