

Simulator Driving Performance Predicts Accident Reports Five Years Later

Lesa Hoffman
University of Nebraska—Lincoln

Joan M. McDowd
University of Kansas Medical Center

L. Hoffman, J. M. McDowd, P. Atchley, and R. A. Dubinsky (2005) reported that visual and attentional impairment (measured by the Useful Field of View test and DriverScan) and performance in a low-fidelity driving simulator did not predict self-reported accidents in the previous 3 years. The present study applied these data to predict accidents occurring within a subsequent 5-year period ($N = 114$ older adults, 75% retention rate). Multivariate path models revealed that accidents in which the driver was at least partially at fault were significantly more likely in persons who had shown impaired simulator performance. These results suggest that even low-fidelity driving simulators may be useful in predicting real-world outcomes.

Keywords: simulator driving, attention, accident risk

One of the unfortunate correlates of increasing age is increased susceptibility to and injury from automobile accidents, particularly in lower mileage drivers. Although there is some debate over the extent to which accident risk increases with age after controlling for mileage driven (e.g., Langford, Methorst, & Hakamies-Blomqvist, 2006), adults over 65 years of age are overrepresented in fatal traffic crashes (Skyving, Hans-Yngve, & Laflamme, 2009). The increasing size of this population due to the aging of the “baby boomer” generation makes older adults’ driving a significant public health issue (Lyman, Ferguson, Braver, & Williams, 2002). Yet given the deleterious effects of driving cessation on mobility and general well-being (e.g., Fonda, Wallace, & Herzog, 2001; Marottoli et al., 2000), paired with empirical data suggesting that blanket age restrictions are unfair and ineffective (e.g., Ball & Owsley, 1993; Langford, Bohensky, Koppel, & Newstead, 2008), it has become increasingly important to find empirically supported ways with which older adults who may be at increased risk for accidents can be distinguished from those older adults who are likely to continue driving safely.

Considerable research efforts have focused on constructing and evaluating measures with which to predict accident risk in older adults. Most promising are test batteries that take into account the multifaceted abilities required for driving safely, including sensory, motor, and cognitive abilities (Anstey, Wood, Lord, & Walker, 2005; Wood, Anstey, Kerr, Lacherez, & Lord, 2008). In evaluations of the contribution of visual attention to the prediction of automobile accidents, one such measure has received more exploration than any other. The Useful Field of View (UFOV) test purports to measure the size of the attentional window through

concurrent central discrimination and peripheral localization tasks that form subtests of processing speed, divided attention, and selective attention. The UFOV test has been shown to significantly predict automobile accidents, predominantly in groups of research participants who have been oversampled for previous accident history or visual impairment (Ball & Owsley, 1993; Owsley, Ball, Sloane, Roenker, & Bruni, 1991). The divided attention subtest in particular has been highlighted as a unique predictor of accidents in larger samples (Ball et al., 2006), although similar results were not found for on-road driving tests (Wood et al., 2008). Not all studies have reported significant accident prediction by the UFOV, however (Brown, Greaney, Mitchel, & Lee, 1993; Hennessey, 1995; Hoffman, McDowd, Atchley, & Dubinsky, 2005).

Other facets of attention have also begun to receive empirical study. Bédard et al. (2006) reported that greater susceptibility to inhibition of return, an attentional mechanism that directs search away from locations or objects already visited, significantly predicted better on-road driving test performance in a group of older adults. Baldock, Mathias, McLean, and Berndt (2007) found that an attention task analogous to the UFOV but also incorporating movement significantly predicted errors by older drivers in an on-road driving task. Finally, Hoffman et al. (2005) described the unique contribution of a change detection measure of attentional search (DriverScan; Hoffman, Yang, Bovaird, & Embretson, 2006) in predicting simulated driving. Results indicated that approximately half of the variance in a latent factor of simulator driving performance was accounted for by measures of visual impairment and visual attention, with significant unique prediction by the UFOV divided attention subtest and the DriverScan test.

Unfortunately, very few studies have examined the extent to which such measures of attention might predict future driving impairment, rather than past or concurrent impairment. Thus, in the present study we aimed to augment this line of work by reporting the findings of a 5-year follow-up to Hoffman et al. (2005), in which we examined the utility of those measures of attention (UFOV, DriverScan) in predicting reports of subsequent accidents in older adults. Although Hoffman et al. (2005) did not find any relationships between attentional ability and reports of

Lesa Hoffman, Department of Psychology, University of Nebraska—Lincoln; Joan M. McDowd, Landon Center on Aging, University of Kansas Medical Center.

We thank Hafía Gregory for her assistance in collecting the follow-up data and the State Farm Companies Foundation for supporting the original data collection.

Correspondence concerning this article should be addressed to Lesa Hoffman, Department of Psychology, 238 Burnett Hall, University of Nebraska—Lincoln, Lincoln, NE 68516. E-mail: lhoffman2@unl.edu

previous accidents, it is an open question as to how attentional ability might predict future accidents instead.

Although real-world accidents are undoubtedly the most relevant outcome, they are generally rare occurrences, and thus statistical power to predict them can be limited. Performance in an on-road driving test may be the next best alternative, but collecting such data can be exceedingly labor intensive and may not be practical for large-scale, real-world assessment. In addition, for safety reasons, the difficulty of the on-road driving task generally must be minimized. In contrast, performance in a driving simulator can be one way in which to obtain measures of driving ability within a safe and controlled environment that are nonetheless capable of demonstrating greater individual variation due to the range of challenges that can be incorporated. Although driving simulators are necessarily artificial environments, research has suggested that performance in a driving simulator does correlate with on-road performance in older adults (e.g., Lee, Cameron, & Lee, 2003) and in beginning drivers as well (e.g., de Winter et al., 2009).

Of interest in the current study is the extent to which previous performance in a driving simulator (as measured in Hoffman et al., 2005) might also predict future accidents. If significant prediction were found, this would suggest that such test batteries (measures of attention, simulated driving performance) might be useful as real-world screening devices in practical settings, in addition to simply serving as measures of individual differences for use in empirical research.

Method

Participants

The Hoffman et al. (2005) original sample was collected in the summer of 2003 and consisted of 152 community-dwelling, currently licensed drivers from a Midwest metropolitan area. It included 68 men (44%) and 87 women (56%) between 63 and 87 years of age ($M = 75.2$ years, $SD = 4.7$). Most were White and had at least a bachelor's degree. We were able to obtain follow-up data from 114 persons (48 men and 66 women; 75% retention), currently between 68 and 90 years of age ($M = 79.9$ years, $SD = 4.7$), from the original sample. Of the 38 participants who did not participate, 14 (7 men and 7 women) declined to participate, 18 (8 men and 10 women) could not be reached, and 6 (5 men and 1 woman) were deceased. Additional analyses comparing participants who were or were not retained for follow-up are reported below.

Measures and Procedure

Participants were contacted in the fall of 2008 and asked whether they would be willing to complete a brief telephone questionnaire about their current driving habits. Only three participants reported that they no longer drove. In addition, participants were asked to provide the approximate dates and circumstances of any incident in which their car struck another car or object or in which they received a traffic ticket in the last 5 years (i.e., since completion of the previous study). Accident responses were reviewed by two experimenters and coded into three types: *no accident* ($N = 91$, 79.8%), *no-fault accident* (e.g., in which participants reported that the actions of someone else were responsible for damage to their car; $N = 9$, 7.9%), and *at least partially-*

at-fault accident (e.g., in which participants reported that their own actions caused damage to their car; $N = 14$, 12.3%). The proportion of agreement in differentiating these two kinds of accidents was 95% (there was only one discrepancy). The no-accident and no-fault accident categories were then combined for analysis, as in Hoffman et al. (2005). In addition, 17 participants reported receiving a traffic ticket, 14 of which were for speeding (12.3%). The report of a speeding ticket was used as a second outcome variable in further analyses.

A predictor reflecting tendency to limit driving during the follow-up period was constructed as a mean of four items assessing the extent to which (on a 5-point scale of *almost never*, *occasionally*, *sometimes*, *very often*, and *almost always*) participants reported avoiding driving in bad weather, at night, or during heavy traffic or avoided busy intersections. The resulting mean was 2.22 ($SD = .96$, $N = 111$), indicating that on average participants occasionally to sometimes limited their driving in these situations. This predictor was centered at 2 for analysis.

Finally, additional model predictors were chosen from variables collected from the 2005 study. Demographic predictors included age (centered at 75 years) and sex (coded 0 = men, 1 = women). Visual impairment was represented by a latent trait estimate ($M = -0.09$, $SD = 0.88$) derived from a measure of static acuity and five levels of contrast sensitivity. Attentional search was represented by a latent trait estimate derived from speed and accuracy of change detection on the DriverScan test (see Hoffman et al., 2006; $M = -0.05$, $SD = 0.98$). The UFOV subtest presentation time thresholds were converted into z scores for ease of interpretation, and each subtest served as a separate predictor (processing speed, $M = -0.07$, $SD = 0.90$; divided attention, $M = -0.02$, $SD = 0.95$; selective attention, $M = 0.05$, $SD = 1.02$). In all predictor variables for vision and attention, higher values indicated greater impairment. Finally, simulator impairment was represented by a latent trait estimate ($M = -0.04$, $SD = 0.79$) derived from simulator outcome measures of lane position variability, proportion of missed divided attention tasks, frequency of crashes, number of stoplight violations, number of speeding violations, and course completion time. Except for simulator driving impairment, which was available for only 97 participants, predictor information was available for all 114 follow-up participants. Additional information about these predictors is available in Hoffman et al. (2005).

Results

Attrition Analyses

No differences were found on the predictor variables from the original study between participants who did or did not complete the follow-up study. Relative to the retention sample of 114 persons, the 38 persons who were not retained did not differ by sex, $\chi^2(1) = 1.27$, $p > .05$; age, $F(1, 150) = 0.86$, $p > .05$; visual impairment, $F(1, 150) = 2.66$, $p > .05$; UFOV processing speed, $F(1, 150) = 2.36$, $p > .05$; UFOV divided attention, $F(1, 150) = 0.28$, $p > .05$; UFOV selective attention, $F(1, 150) = 1.61$, $p > .05$; DriverScan attentional search, $F(1, 150) = 0.68$, $p > .05$; or simulator driving impairment, $F(1, 150) = 0.83$, $p > .05$.

Path Analyses

Bivariate correlations between the predictors and outcomes within the retention sample ($N = 114$) are given in Table 1. Because of the restricted range of the binary outcomes, we obtained correlations within Mplus Version 5.2 using weighted least squares parameter estimates with a mean-adjusted and variance-adjusted chi-square test statistic (Muthén & Muthén, 2007). Pearson correlations were estimated between continuous variables, biserial correlations were estimated between continuous and binary variables, and tetrachoric correlations were estimated between the binary variables. None of the bivariate relationships with accident reports reached significance. Reports of speeding tickets were significantly less likely in women and in persons with greater impairment in UFOV divided and selective attention. Tendency to limit driving was significantly higher in women and in persons with greater impairment in UFOV divided attention and was marginally higher in persons with greater impairment in UFOV processing speed and DriverScan attentional search.

We then examined these relationships in a multivariable and multivariate path model estimated via full information maximum likelihood with Monte Carlo integration using Mplus Version 5.2, in which logit links were specified for the accident and speeding ticket binary outcomes. As shown in Figure 1, the model was just-identified, such that all possible unique paths were examined, and convergence was obtained with no issues. Unstandardized coefficients on the logit metric (i.e., the effect of a one-unit change in the predictor on the log odds of the outcome) and their standard errors from this model are provided in Figure 1; odds ratios and standardized coefficients are provided below for marginal or significant effects. As shown on the left-hand side of Figure 1 (top estimates), and as reported in the Hoffman et al. (2005) study, simulator impairment was significantly related to impairment in DriverScan attentional search ($Est_{std} = 0.27$) and marginally related to impairment in UFOV divided attention ($Est_{std} = 0.20$). Also, as shown on the left-hand side of Figure 1 (bottom estimates), no significant unique predictors of tendency to limit driving were found.

The estimated model threshold for the binary accident outcome on the logit (log odds) metric was 2.05 ($SE = 0.53$), such that at the reference point of all predictors (i.e., men age 75 who occa-

sionally limit their driving and who had average scores on the other predictors), 89% of the sample was predicted to have reported no accidents (or a no-fault accident). With respect to the unique effects of the predictors, the log odds of reporting an at least partially at fault accident were significantly higher for those who had greater impairment in the driving simulator ($Est_{logit} = 1.45$, odds ratio = 4.25, $Est_{std} = 0.50$). No other significant unique effects were found.

The estimated threshold for the binary speeding ticket outcome on the logit (log odds) metric was 1.72 ($SE = 0.50$), such that at the reference point of all predictors (i.e., men age 75 who occasionally limit their driving and who had average scores on the other predictors), 85% of the sample was predicted to have reported no speeding tickets. With respect to the predictors, the log odds of receiving a speeding ticket were significantly lower for women ($Est_{logit} = -1.38$, odds ratio = 0.25, $Est_{std} = -0.29$) and for those with greater impairment on the UFOV selective attention task ($Est_{logit} = -1.14$, odds ratio = 0.32, $Est_{std} = -0.48$). The log odds of reporting a speeding ticket were actually marginally higher for those with greater impairment on DriverScan ($Est_{logit} = 0.96$, odds ratio = 2.60, $Est_{std} = 0.39$). No other significant unique effects were found.

Given the relatively small sample size and the complexity of model estimated, the possibility of null results due to lack of statistical power was investigated via Monte Carlo power simulations in which the obtained model estimates were used as population values and 1,000 replications were conducted. In this approach, the proportion of replications in which a given estimate is significant provides an empirical estimate of power. Only previous simulator performance (the only significant predictor) obtained a power estimate of 80% or more in detecting follow-up accidents. Power to detect most other paths was very low, which is not surprising given the very small effects found (standardized path coefficients ranging from $r = \pm .04$ to $\pm .19$). The exception to this pattern was UFOV divided attention, which had an empirical power rate of 62% to detect a standardized effect size of $r = -.34$ ($p = .107$; i.e., a relationship such that greater attentional skill was actually related to a higher likelihood of having an accident). The same pattern of lower power due to very small effects was ob-

Table 1
Bivariate Correlations Between Predictor and Outcome Variables

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|-----------------------------------|------|-------|------|------|------|------|------|------|------|------|----|
| Outcomes | | | | | | | | | | | |
| 1. Accident report | — | | | | | | | | | | |
| 2. Speeding ticket report | -.14 | — | | | | | | | | | |
| Predictors | | | | | | | | | | | |
| 3. Reported limited driving | -.25 | -.22 | — | | | | | | | | |
| 4. Simulator impairment | .21 | -.22 | .14 | — | | | | | | | |
| 5. Sex (0 = men, 1 = women) | -.23 | -.35* | .21* | .13 | — | | | | | | |
| 6. Original age | .04 | -.20 | .05 | .16 | -.03 | — | | | | | |
| 7. Visual impairment | -.00 | -.18 | .02 | .05 | -.05 | .22* | — | | | | |
| 8. UFOV processing speed | -.15 | -.28 | .12 | .17* | -.10 | .12 | .13 | — | | | |
| 9. UFOV divided attention | -.24 | -.31* | .17* | .40* | .04 | .25* | .17 | .30* | — | | |
| 10. UFOV selective attention | -.08 | -.43* | .03 | .38* | -.09 | .36* | .29* | .29* | .52* | — | |
| 11. DriverScan attentional search | -.16 | -.14 | .15 | .43* | .06 | .41* | .20* | .20* | .45* | .60* | — |

Note. UFOV = Useful Field of View test.
* $p < .05$.

search task or the UFOV tasks. One possible explanation for this null result is simple range restriction: Only 12% of the retention sample reported an at least partially at fault accident, and most of these reports described relatively minor occurrences (e.g., backing out of a parking space into another car). Previous studies that have found significant prediction of accidents by the UFOV test generally have either deliberately oversampled persons with visual impairment and history of previous accidents (Ball & Owsley, 1993; Owsley et al., 1991) or examined accident reports in much larger samples (Ball et al., 2006). Another possible explanation concerns systematic differences in underreporting: If persons with lower cognitive ability were less inclined to report such detrimental driving incidents, or perhaps less likely to remember them, the expected deleterious effect of attentional impairment on driving would be lessened, at the very least. We note, in potential support of this idea, that the observed predictive effects of the attention measures were actually negative, or backwards, such that persons with greater impairment were (nonsignificantly) less likely to report an accident. The largest nonsignificant but negative effect was found for impairment in the UFOV divided attention task, a measure that had been singled out in other studies as a positive predictor of driving impairment instead.

Finally, we note the observed effects for reported speeding tickets, which were significantly less likely in women and in persons with greater impairment in UFOV selective attention but marginally more likely in persons with greater impairment in DriverScan attentional search. The results are contradictory, perhaps in part due to the multifaceted causes of receiving a speed ticket itself. Although speeding tickets could be viewed as a negative outcome if they occurred due to inattention (e.g., not realizing one had drifted above the posted speed limit), they might also be viewed as a positive outcome if they reflected confidence or greater ability, such that those drivers may still feel more comfortable traveling at higher speeds. Further investigation is needed to distinguish these possibilities and likely many others.

In summary, the present study demonstrated that performance in a low-level driving simulator could significantly predict self-reported automobile accidents 5 years later in a sample of 114 older adults, lending evidence of external validity of such measures. These findings will, it is hoped, augment ongoing efforts at finding ways to identify and predict those individual differences characteristics that differentiate safe older adult drivers from those who may need additional monitoring or testing to ensure their safety and the safety of others.

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 Review, 25*, 45–65. doi:10.1016/j.cpr.2004.07.008
- Arthur, W., Bell, S. T., & Edwards, B. D. (2005). Convergence of self-report and archival crash involvement data: A two-year longitudinal follow-up. *Human Factors, 47*, 303–313. doi:10.1518/0018720054679416
- Baldock, M. R. J., Mathias, J., McLean, J., & Berndt, A. (2007). Visual attention as a predictor of on-road driving performance of older drivers. *Australian Journal of Psychology, 59*, 159–168.
- Ball, K. K., & Owsley, C. (1993). The useful field of view test: A new technique for evaluating age-related declines in visual function. *Journal of the American Ophthalmological Association, 63*, 71–79.
- Ball, K. K., Roenker, D. L., Wadley, V. G., Edwards, J. D., Roth, D. L., McGwin, G., Jr., . . . Dube, T. (2006). Can high-risk older drivers be identified through performance-based measures in a department of motor vehicles setting? *Journal of the American Geriatric Society, 54*, 77–84. doi:10.1111/j.1532-5415.2005.00568.x
- Bédard, M., Leonard, E., McAuliffe, J., Weaver, B., Gibbons, C., & Dubois, S. (2006). Visual attention and older drivers: The contribution of inhibition of return to safe driving. *Experimental Aging Research, 32*, 119–135. doi:10.1080/03610730500511918
- Brown, J., Greaney, K., Mitchel, J., & Lee, W. E. (1993). *Predicting accidents and insurance claims among older drivers*. Hartford, CT: ITT Hartford Insurance Group & American Association of Retired Persons.
- de Winter, J. C. F., de Groot, S., Mulder, M., Wieringa, P. A., Dankelman, J., & Mulder, J. A. (2009). Relationships between driving simulator performance and driving test results. *Ergonomics, 52*, 137–153. doi:10.1080/00140130802277521
- Fonda, S. J., Wallace, R. B., & Herzog, A. R. (2001). Changes in driving patterns and worsening depressive symptoms among older adults. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 56*, S343–S351.
- Hennessey, D. F. (1995). *Vision testing of renewal applicants: Crashes predicted when compensation is inadequate* (Report No. RSS-95-152). Sacramento, CA: California Department of Motor Vehicles.
- Hoffman, L., McDowd, J. M., Atchley, P., & Dubinsky R. A. (2005). The role of visual attention in predicting driving impairment in older adults. *Psychology and Aging, 20*, 610–622. doi:10.1037/0882-7974.20.4.610
- Hoffman, L., Yang, X., Bovaird, J. A., & Embretson, S. E. (2006). Measuring attention in older adults: Development and psychometric evaluation of DriverScan. *Educational and Psychological Measurement, 66*, 984–1000. doi:10.1177/0013164406288170
- Langford, J., Bohensky, M., Koppel, S., & Newstead, S. (2008). Do age-based mandatory assessments reduce older drivers' risk to other road users? *Accident Analysis and Prevention, 40*, 1913–1918. doi:10.1016/j.aap.2008.08.010
- Langford, J., Methorst, R., & Hakamies-Blomqvist, L. (2006). Older drivers do not have a high crash risk—A replication of low mileage bias. *Accident Analysis and Prevention, 38*, 574–578. doi:10.1016/j.aap.2005.12.002
- Lee, H. C., Cameron, D., & Lee, A. H. (2003). Assessing the driving performance of older adult drivers: On-road versus simulated driving. *Accident Analysis and Prevention, 35*, 797–803. doi:10.1016/S0001-4575(02)00083-0
- Lyman, S., Ferguson, S. A., Braver, E. R., & Williams, A. F. (2002). Older driver involvements in police reported crashes and fatal crashes: Trends and projections. *Injury Prevention, 8*, 116–120. doi:10.1136/ip.8.2.116
- Marottoli, R. A., Mendes de Leon, C. F., Glass, T. A., Williams, C. S., Cooney, L. M., Jr., & Berkman, L. F. (2000). Consequences of driving cessation: Decreased out-of-home activity levels. *Journal of Gerontology, Series B: Psychological Sciences and Social Sciences, 55*, 334–340.
- Muthén, L. K., & Muthén, B. O. (2007). *Mplus user's guide* (5th ed.). Los Angeles, CA: Author.
- Owsley, C., Ball, K. K., Sloane, M. E., Roenker, D. L., & Bruni, J. R. (1991). Visual/cognitive correlates of vehicle accidents in older drivers. *Psychology and Aging, 6*, 403–415. doi:10.1037/0882-7974.6.3.403
- Skyving, M., Hans-Yngve, B., & Laflamme, L. (2009). A pattern analysis of traffic crashes fatal to older drivers. *Accident Analysis and Prevention, 41*, 253–258. doi:10.1016/j.aap.2008.11.008
- Wood, J. M., Anstey, K. J., Kerr, G. K., Lacherez, P. F., & Lord, S. (2008). A multidomain approach for predicting older driver safety under in-traffic road conditions. *Journal of the American Geriatric Society, 56*, 986–993. doi:10.1111/j.1532-5415.2008.01709.x

Received October 2, 2009

Revision received February 1, 2010

Accepted February 1, 2010 ■