Visual Attention Problems as a Predictor of Vehicle Crashes in Older Drivers

Karlene Ball,* Cynthia Owsley,† Michael E. Sloane,‡ Daniel L. Roenker,* and John R. Bruni*

Purpose. To identify visual factors that are significantly associated with increased vehicle crashes in older drivers.

Methods. Several aspects of vision and visual information processing were assessed in 294 drivers aged 55 to 90 years. The sample was stratified with respect to age and crash frequency during the 5-year period before the test date. Variables assessed included eye health status, visual sensory function, the size of the useful field of view, and cognitive status. Crash data were obtained from state records.

Results. The size of the useful field of view, a test of visual attention, had high sensitivity (89%) and specificity (81%) in predicting which older drivers had a history of crash problems. This level of predictability is unprecedented in research on crash risk in older drivers. Older adults with substantial shrinkage in the useful field of view were six times more likely to have incurred one or more crashes in the previous 5-year period. Eye health status, visual sensory function, cognitive status, and chronological age were significantly correlated with crashes, but were relatively poor at discriminating between crash-involved versus crash-free drivers.

Conclusions. This study suggests that policies that restrict driving privileges based solely on age or on common stereotypes of age-related declines in vision and cognition are scientifically unfounded. With the identification of a visual attention measure highly predictive of crash problems in the elderly, this study points to a way in which the suitability of licensure in the older adult population could be based on objective, performance-based criteria. Invest Ophthalmol Vis Sci. 1993;34:3110–3123.

The elderly represent the most rapidly growing segment of the driving population in industrialized societies, both in the total number of drivers on the road, and the number of miles driven annually per driver. It is estimated that by the year 2024 one out of four drivers in the United States and Western Europe will be older than 65 years. Older drivers as a group have more traffic convictions and crashes and incur more fatalities per mile driven than any other adult age group. Although the stereotype of the impaired older driver may be true for some persons, it is also the case that a significant number of older adults have excellent driving skills. Older adults exhibit marked individual differences in many skills, and driving is unlikely to be an exception. Little is known, however, about what behavioral and physiological changes in the aging process relate to a decline in driving ability. Given that the personal automobile has become the preferred mode of travel in industrialized societies, and given that the elderly rely on the automobile to maintain mobility, there is a pressing need for research to identify the factors that place certain older drivers at risk for crash involvement.

Driving is a highly visual task, and thus it might be expected that the higher incidence of visual problems and eye disease in the elderly is a primary cause of their driving difficulty. This expectation is reflected by the practice of assessing visual acuity, and sometimes peripheral vision, at driver licensing sites in each state. Despite the intuition that vision and driving ability...
should be related, earlier studies have found only weak correlations between visual deficits (eg, static and dynamic acuity, disability glare) and vehicle crashes.5-7 These weak correlations were often statistically significant due to very large sample sizes, but accounted for less than 5% of the crash variance, and thus are insignificant from the practical standpoint of identifying what older drivers are at risk for crash involvement.

This failure to find a strong link between visual deficits and driving in previous work may be attributable to several factors.2,8 In subject samples from earlier work, there was a preponderance of drivers with zero crashes on record, making it difficult to evaluate a model designed to predict crash frequency. In addition, crashes are rare occurrences and therefore the researcher has the statistical burden of predicting an improbable event. Another reason for weak links in earlier work is that poor vision may cause drivers to modify their behavior, such as avoiding challenging roadway conditions,9 thus reducing their crash risk. Such self-imposed changes in driving behavior may mitigate against a correlation between poor vision and crash involvement. Finally, previous studies relied almost exclusively on visual sensory tests as the independent (predictor) variables, ignoring higher-order perceptual and cognitive components5-7,10 or vice versa.11-13 Sensory tests, such as visual acuity and peripheral field sensitivity, although quite appropriate for the clinical diagnosis and assessment of ocular disease and vision loss, do not by themselves reflect the visual complexity of the driving task, and therefore would not be expected to reveal a strong relationship between vision and driving. The visual demands of driving are quite intricate. Controlling a vehicle takes place in a visually cluttered environment and involves the simultaneous use of central and peripheral vision and the execution of both primary and secondary visual tasks. The driver is usually uncertain as to when and where an important visual event will occur. Visual sensory tests do not typically incorporate these stimulus and task features, but instead seek to minimize perceptual and cognitive influences. Thus, in designing a visual test to predict driving problems, it would be advantageous to incorporate stimulus and task characteristics that more accurately reflect the visual demands of driving.14

For these reasons an approach limited to visual sensory factors is by itself inadequate in identifying factors that place older adults at risk for driving problems. A promising alternative is to also examine the role of higher-order attentional factors in driving by the elderly. Our previous studies have indicated that some older adults have an impairment in the "preattentive level" of visual attention.15,16 Other aspects of attention, such as attention switching and serial visual search, can also be impaired in older adults.17,18 Visual attention is used to direct information processing resources to potentially important visual events, a critical skill for avoiding crashes. Thus, impaired attentional skills could be a significant feature that distinguishes crash-involved older drivers from those who are crash-free. Some earlier work is consistent with this hypothesis. Analyses of police accident reports indicate that most crashes by older drivers are caused by alleged "driver inattention."19 In addition, several studies of commercial drivers found an association between selective attention problems and increased amount of crashes.20-22

Recently we developed a regression model for predicting crash frequency in elderly drivers on the basis of a preliminary study that assessed visual and cognitive skills in a small sample of older (57 to 83 years) adults.9 Figure 1 portrays this model. The most prominent feature of the model is that visual attention and mental status are the only variables that significantly predict crash frequency. Although the model acknowledges that eye health is related to visual sensory function, and visual sensory function is related to visual attention, neither eye health nor visual sensory function is related to crashes. This initial study was critical in developing the model, however, it was far from conclusive because most of our independent variables were restricted in range and our analysis was limited to a total of only 25 crashes. Thus the large sample study described here was deemed crucial for evaluating this model.

![Multiple Regression Model for Predicting Crash Frequency](image)

**FIGURE 1.** Multiple regression model for predicting crash frequency in elderly drivers. This model was developed on the basis of a small-sample, retrospective study from our laboratory, as described in detail in our earlier report.9 The only independent variables in this model that significantly predict crash frequency during the previous 5-year period are UFOV (a measure of visual attention) and mental status. Eye health is significantly related to visual function, which in turn is related to UFOV; however, eye health and visual function do not further improve the prediction of crash frequency. This model is evaluated by the present large-sample study, which avoided several of the methodological limitations of the earlier work.9
The test of visual attention employed in the earlier study consisted of a central target identification task coupled with a peripheral target localization task, which together provided a measure of the size of the useful field of view (UFOV).\textsuperscript{15,16} The UFOV provides a measure of the spatial area within which a person can be alerted to visual stimuli in a variety of situations;\textsuperscript{23-25} it is conceptually distinct from the visual sensory field that describes luminance sensitivity throughout the field.\textsuperscript{26} For example, preliminary work indicated that for all older observers without significant visual field loss, the peripheral localization task could be performed with 100\% accuracy at all eccentricities evaluated (10°, 20°, and 30° visual angle) without the central task and distractors. However, some older observers required longer stimulus durations than others to achieve this level of performance, which indicates a relatively slower speed of visual processing. With the addition of the central task, the localization performance of a subset of these observers was hampered, such that the most eccentric targets could no longer be localized (a reduction in the size of the UFOV). This reduction could, however, be counteracted by increasing stimulus duration to regain 100\% accuracy in the dual task situation. Finally, the addition of distractors also impaired localization performance for a subset of older observers, again such that the most eccentric targets could no longer be localized. This effect could sometimes be reversed by increasing target duration to compensate for the distractor effect. The size of the UFOV, defined as that eccentricity at which observers can localize peripheral targets correctly 50\% of the time, is thus a dynamic measure. It is a function of (at least) three test variables: the duration of target presentation, the level of complexity of a secondary central task, and the salience of a peripheral target (where no distractors represents the optimal salience). Thus, the UFOV test incorporates stimulus and task features that seem critical for driving. The current study uses a large sample of older drivers to test this model, assessing various aspects of visual information processing including health status of the visual system, visual sensory function, visual attentional skills, and cognitive skills.

MATERIALS AND METHODS

The recruitment population consisted of all licensed drivers aged 55 years and older who lived in Jefferson County, Alabama (\(N = 118,553\)). Ultimately, we wanted to achieve a stratified sample that was balanced with respect to two variables, crash frequency during the previous 5-year period, and age. We defined three categories of crash frequency for the previous 5-year period (0, 1 to 3, and 4 or more) and seven age categories (55 to 59, 60 to 64, 65 to 69, 70 to 74, 75 to 79, 80 to 84, and 85+ years). Crash data were provided by the Alabama Department of Public Safety, the state agency that compiles records on all drivers licensed by Alabama. Sampling was designed to achieve an equal number of drivers in each of these 21 cells rather than to represent the true percentages of these groupings in the actual driving population. It was important that the sample include a wide range of crash frequencies to avoid a restriction of range on the dependent variable. A sampling strategy that reflected actual crash prevalence in the general population would result in a sample primarily consisting of zero-crash drivers, which would be counterproductive in our effort to evaluate the model. Thus our sampling strategy was designed to over-represent (in terms of the general population) those persons with multiple crashes because weighting factors in the final analyses can be used to simulate the true population. It was also important to include a wide range of ages because it would allow determination of whether relationships between the independent variables and crashes vary with age. Our total population of older drivers was first sorted into these 21 cells, and 75 drivers were randomly selected from each cell. Contact letters were sent to those persons who were listed in the local telephone directory (\(N = 1,342\)). The letter stated that the study was about how vision problems in older adults affected their visual activities of everyday life, but it did not specifically mention driving because the overall study dealt with many kinds of activities, not just driving.

A total of 302 subjects were tested. Six of these participants were later excluded from the sample because they had not driven in the previous 5 years, although they maintained a current driver's license. Two additional participants were excluded because they did not finish the test protocol. The final sample had 294 participants, with 33\% of subjects having 0 crashes, 49\% with 1 to 5 crashes, and 18\% with 4 or more crashes. Within each crash frequency category, age was evenly distributed and thus met the requirements of our sampling strategy. The mean age of the entire sample was 71 years (range, 56 to 90 years); 136 were men and 158 were women. All participants lived independently in the community. Before participation, written informed consent was obtained from each subject after the nature of the study and contents of the protocol were explained. The tenets of the Declaration of Helsinki were followed, and approval of the study was granted by the Institutional Review Board for Human Use of the University of Alabama at Birmingham.

Although all subjects were legally licensed to drive, this does not mean that they met the minimum
Acuity is assessed in this state only at the time one initially applies for a license (typically in young adulthood), or on rare occasions, when a driver is referred by someone (e.g., physician, police officer) who believes that the driver is a threat to himself or other road users. Many states (74%) have some form of a routine retesting program where the licensee must undergo reexamination (e.g., vision test, road test, knowledge test). But this is not the case in Alabama, and thus our use of licensed drivers does not necessarily mean that our sample’s range on the acuity variable is truncated at a specific acuity value (see Figure 2 in Results).

There were five parts to the protocol, which were completed in a single visit to the laboratory: visual sensory function, mental status, UFOV, driving habits questionnaire, and eye health. The order of the five parts was counterbalanced across subjects, except for the eye health examination, which was always last. The visual sensory function tests consisted of visual acuity, contrast sensitivity, disability glare, stereopsis, color discrimination, and visual field sensitivity. These tests were described in detail in our earlier article, and are summarized here. Visual acuity was measured with the Bailey-Lovie chart, and expressed as log minimum angle resolvable. Contrast sensitivity was measured with the Pelli-Robson Contrast Sensitivity Chart, and expressed as log contrast sensitivity. Disability glare was measured with the MCT-8000 (Vistech Consultants, Dayton, OH), and defined as the difference in letter acuity (log minimum angle resolvable) under conditions of glare versus no glare. Stereopsis was measured with three clinical tests (Randot, TNO, and Frisby). Color discrimination was measured with the extended D-15 test. Color field sensitivity was measured with the Humphrey Field Analyzer using the screening program for the central 60° with the quantify defects option. In our use of this program we used a preset initialization value of 34 dB (both central and peripheral), which serves as a baseline or normal visual field against which performance is compared. The standard protocols for all these tests were followed, as described in the manufacturers’ directions.

All tests were binocular except the visual field test in which each eye was tested separately (Humphrey Program 30-2, 0 to 60°). For most tests subjects wore their own habitual optical correction because their everyday visual performance was of interest. However, if a test of visual function specifically called for a near-correction in the standard instructions for administering the test, we followed those instructions (e.g., Humphrey Field Analyzer, Pelli-Robson chart). These specific tests were chosen because they represent major aspects of visual sensory function and have good test-retest reliability. Although several of these visual function measures were unrelated to crash frequency in our earlier study, they were still included in the current study because of restriction of range problems on these variables in the earlier study.

Mental status was assessed by the Mattis Organic Mental Status Syndrome Examination, specifically designed to assess cognitive status in the elderly. This test provides a composite score of cognitive function that reflects performance in several categories such as abstraction, digit span, verbal and visual memory, and block design. Additional cognitive tests were carried out to evaluate visuospatial abilities and included the Rey-Osterreith test, the Trailmaking test, and the block design of the Wechsler Adult Intelligence Scale (Revised).

The size of the UFOV was assessed using the UFOV Visual Attention Analyzer, Model 2000 (Visual Resources, Inc., Bowling Green, KY). This microprocessor-based instrument uses three subtests that provide a reliable measure of UFOV size, expressed in terms of the percentage of reduction (0 to 90%) of a maximum 35° radius field. A similar version of the procedure has been described in detail elsewhere. Briefly, in the first subtest, which was designed to assess speed of visual processing, the subject was required to identify a target of varying duration, presented in the fixation box. This target was the silhouette of a car or a truck. The second subtest, designed to assess the ability to divide attention, also required the identification of the central target, in addition to the localization of a simultaneously presented peripheral target (a silhouette of a car). This target appeared unpredictably at each of 24 different peripheral locations along 8 radial spokes (4 cardinal and 4 oblique) at 3 eccentricities (targets centered at 10°, 20°, or 30°). The duration of the display was varied to measure speed of visual processing for this divided attention task. The third subtest, designed to assess selective attention abilities, required these same two responses (also at different stimulus durations), however, the peripheral target was embedded in distractors (triangles). This subtest provided yet another measure of speed of processing when distracting stimuli were present.

These subtests were presented on a 20-inch (diagonal) video monitor at a viewing distance of 23.5 cm. Targets were presented at high contrast (99%), and subtended 5.1 horizontal × 3.2 vertical degrees of visual angle. Previous work that investigated the effect...
FIGURE 2. Frequency distributions for each independent variable in the model (i.e., number of subjects for different values of each independent variable).
of poor visual acuity on UFOV performance demonstrated that blur, even with the addition of +5 diopters, had no effect on radial localization of peripheral targets with or without distractors. Thus observers wore their current optical correction during testing. Observers indicated their identification of the centrally presented target by touching the image they had seen in the fixation box in response to a prompt on the screen after each trial. Peripheral localization responses were similarly made by touching the spoke on the screen on which they perceived the peripheral target to have been presented. For subtest 1, the minimum duration at which subjects could perform the task with 75% correct was noted. For subtests 2 and 3, the best fitting line reflecting the relationship between eccentricity and localization errors was first computed for each test duration, and the size of the UFOV was defined for that stimulus duration as that eccentricity at which the subject could localize the peripheral target correctly 50% of the time. Performance in each of the three subtests was then scaled, in each case along a duration continuum, to arrive at three scores representing the extent of difficulty with respect to speed of processing, divided attention, and selective attention. These scores ranged from 0 (no problem) to 30 (great difficulty). For example, within subtest 1, if an observer was unable to identify the central target correctly 75% of the time at the longest stimulus duration (240 ms), that observer received a score of 30 (maximum deficit) on subtest 1. If, however, this 75% criterion was achieved at an average of 40 ms or less then that observer received a score of 0 (no deficit) on subtest 1. Similarly, in subtests 2 and 3, if the computed size of the UFOV was less than 5° at 240 ms then a score of 30 (maximum deficit) was assigned. However, if the computed size of the UFOV was 30° at a duration of 40 ms then a score of 0 (no deficit) was assigned. A previous study comparing the size of the UFOV among a large number of persons exhibiting distractor and/or divided attention and/or slowing problems relative to a group of persons exhibiting none of these problems found that deficits in each of these abilities were both independent and additive in their effect on the size of the UFOV. For example, those persons with only one of the three problems experienced a reduction in UFOV size of 30% to 45%, whereas those with all three problems demonstrated a 90% restriction in their UFOV relative to those with none of these problems. Therefore, to summarize UFOV performance, the three scale scores were summed to yield a composite score between 0 and 90, which represented the total percentage reduction of the UFOV.

All subjects received a detailed eye health examination by an ophthalmologist, which included direct and indirect ophthalmoscopy after dilation, biomicroscopy, applanation tonometry, a refraction for distance, and an assessment of external eye health. A three-point rating scale, as described in our earlier study, was used to determine to what extent clinical changes in the eye would be expected to cause a functional problem in each of three broad categories: central vision problem, peripheral vision problem, and ocular media problem. In addition, each subject was assigned to a primary diagnostic category (eg, normal, cataract, macular disease).

We administered a questionnaire about the subject’s driving habits such as: (1) driving exposure (eg, how many miles/year, how many days/week, how many trips/day); (2) avoidance of potentially challenging driving situations (eg, left turns across traffic, driving alone); (3) number of crashes incurred during the previous 5-year period in which the police came to the scene. In addition to this “self-report” crash information, crash frequency during the previous 5-year period was obtained for each subject from the state computer of the Alabama Department of Public Safety, as discussed earlier. After data collection was completed, the written accident reports (filed by the officer at the scene) for all subjects were obtained from the state, which detailed the circumstances surrounding each crash.

**DEFINING THE DEPENDENT VARIABLE**

The dependent variable used in this study was the total at-fault crashes recorded by the state during the 5-year period before testing. There are two reasons for choosing state recorded crash data, rather than that which is self-reported. First, in our previous study, self-reported crash frequency information was unreliable for older drivers when compared to state records. In an effort to improve the validity of the self-report information, we increased the number of items concerned with crash involvement in our questionnaire. However, in the current sample as well, there was not good agreement between the number of state-recorded and self-reported crashes (eg, only 27% of those drivers with at least one crash reported the same number of crashes as recorded by the state), and the lack of agreement was primarily reflected in a lower frequency of self-report than crashes listed on the state record. A second reason for choosing state-recorded crash frequency as the dependent measure is that accident reports from the state provide detailed information about the circumstances surrounding each crash, which is useful for determining fault and for subdividing crashes into different types. Three raters independently studied each accident report to determine whether our driver was at fault. Two of the raters did not know which of the drivers referred to in the accident report was the research participant. De-
tained examination of the 559 accident reports revealed 195 crashes where our research participant was clearly not at fault (eg, subject’s unoccupied parked vehicle was hit). Concordance among the three raters was perfect in identifying these cases. In the remaining 364 accident reports our research participant was judged to be solely or partially at fault. Although concordance was not perfect (83%) regarding the degree of fault involved in these crashes, all three raters always agreed that our subject was at least partially to blame for the crash. Thus our sample of 294 participants were involved in 364 at-fault crashes.

Some mention should also be made about the length of time over which crash data were cumulated (5 years). This period of time was chosen for statistical reasons. Total “at-fault” crashes were initially sorted into 1-year groupings, depending on when a crash occurred relative to the date of testing. Crashes were then cumulated over varying amounts of time to obtain several crash frequencies per participant (1 year, 2 years, 3 years, 4 years, and all 5 years previous to testing). Due to the fact that crashes are rare events, sampling crashes over a shorter duration (eg, 1 year) produced a severe restriction of range on the dependent variable such that zero-order correlations of all predictors with crash frequency were greatly attenuated relative to a longer period of sampling. A graph relating the number of years over which crashes were cumulated to the strength of the correlations between crash frequency and the independent variables revealed that the 5-year sample maximized the predictive relationships of all variables. Thus the longest sampling period of 5 years was chosen for the development of a predictive model.

Previous work on driver safety has sometimes used crash rate (number of crashes/number of miles driven during given period of time), rather than crash frequency (number of crashes during given period of time), as the dependent variable. In the calculation of rate, the number of miles driven, or driver mileage, over some period of time is typically determined by the subject’s own estimates. However, for the following reason we chose crash frequency, not rate, as the dependent measure in this study. To obtain an estimate of driver mileage for each subject, we included several different types of questions on the driving habits questionnaire about driving exposure; these questions were designed to converge on a single mileage figure per subject. Subjects were asked about mileage from a number of perspectives, such as how many miles were driven during a week’s time, how many trips taken per day, and the average distance of a trip. However, the responses to these questions yielded a group of inconsistent estimates of number of miles driven and thus the mileage correction was not performed.

**RESULTS**

The goal of this study was to test a model designed to predict crash frequency in older drivers on the basis of visual and cognitive measures. Before evaluating this model, the correlation matrix among all independent variables and the dependent variable was computed. This matrix made two things apparent. First, variables within a given aspect (eye health, visual sensory function, visual attention, mental status) of visual information processing were highly interrelated. Second, variables between different aspects of visual information processing were moderately related. To reduce the redundancy among our independent measures within each aspect of visual processing evaluated (eg, problem of multiple collinearity), we chose the variable with the highest zero-order correlation to crash frequency, to represent its class of variables within the model. With respect to the UFOV, a preliminary analysis of each of the subset scale scores relative to the other measures in the model indicated that the composite score was by far the strongest correlate of crash frequency (the strongest individual subset score was from task three, r = 0.31). This procedure resulted in this list of independent variables to be further evaluated: (1) A measure of eye health status, which was the sum of all three types of ratings as described earlier; (2) Two measures of visual sensory function, one representing central vision (contrast sensitivity), and the other representing peripheral vision (sensitivity loss in the 30 to 60° region of the visual field, averaged across both eyes and all meridians). Because functional status in central and peripheral vision are dissociated in many ocular diseases, both types of measures were included as independent variables.*; (3) Composite measure of the UFOV, expressed in terms of percentage of reduction, which represented visual attention; and (4) The Mattis Organic Mental Status Syndrome Examination composite, representing mental status.

The diagnostic category for ocular disease was not related to crash frequency, and thus was not used in the model. The diagnostic breakdown of the sample was: 135 with normal eye health, 100 cataract, 23 retinal disease, 6 glaucoma/ocular hypertension, 5 diabetic retinopathy, and 26 miscellaneous. To more fully evaluate the relationship between eye health diagnosis per se and crash frequency, a large sample would be required in each diagnostic category to be evaluated. However, it is doubtful that diagnosis by itself will be identified as a potent predictor of crashes, given that*

*Although letter acuity is the preferred method of assessing visual function at licensing sites, we did not use it as the measure representing central vision in our model, for the following reason. Crash frequency was more strongly correlated with contrast sensitivity (r = 0.24) than was acuity (r = 0.225). Thus, contrast sensitivity was chosen to represent central vision in the model.
Visual Factors and Vehicle Crashes in the Elderly

TABLE 1. Correlations Among Variables in the Model*

<table>
<thead>
<tr>
<th>Eye Health</th>
<th>Central Vision</th>
<th>Peripheral Vision</th>
<th>Mental Status</th>
<th>UFOV</th>
<th>Crash Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eye Health</td>
<td>1.0</td>
<td>−.67</td>
<td>.50</td>
<td>.24</td>
<td>.40</td>
</tr>
<tr>
<td>Central Vision</td>
<td>1.0</td>
<td>−.57</td>
<td>−.20</td>
<td>−.47</td>
<td>−.24</td>
</tr>
<tr>
<td>Peripheral Vision</td>
<td>1.0</td>
<td>.29</td>
<td>.48</td>
<td></td>
<td>.26</td>
</tr>
<tr>
<td>Mental Status</td>
<td>1.0</td>
<td>.48</td>
<td>.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UFOV</td>
<td>1.0</td>
<td>.52</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* All coefficients were statistically significant, \( P < 0.01 \), \( n = 294 \). The strongest correlate of crash frequency during the prior 5-year period was percentage of UFOV reduction. (The correlation between central vision and the other variables is negative because central vision is expressed in terms of log contrast sensitivity; that is, higher numbers represent better performance. For the other independent variables, lower numbers represent better performance.)

the functional manifestation of an ocular condition and the ability to compensate for vision impairment varies widely across patients.

All the independent variables used in the model had similar positively skewed distributions, as illustrated in Figure 2. In addition, all variables covered their entire possible range. Significant numbers of subjects (over 10% of the sample) were represented at the “poorer” end of the continuum on each variable, thus ruling out a restriction of range problem that would affect the predictive ability of any variable in the model. As shown in Table 1, all variables in the model were significantly related to crash frequency. UFOV was the strongest correlate. Eye health, central and peripheral vision, and mental status were more highly related to UFOV than to crash frequency per se. Using these measures from different aspects of the visual information processing system (as listed in Table 1), we tested our original model using the LISREL VII structural modeling program. This program analyzes the covariance matrix among the variables to arrive at a system of simultaneous linear equations that allows the dependent variable (crash frequency) to be expressed in terms of the structural relationships among the independent variables. Of special relevance to our goal, LISREL enables one to evaluate independent variables in terms of whether they directly influence the dependent variable, or whether they operate indirectly through other variables, and thus LISREL is superior to a multiple regression approach for model development. As shown in Figure 3, our model as formulated postulates that eye health, central vision, and peripheral vision have only indirect effects on crash frequency but direct effects on visual attention (UFOV). It further asserts that mental status has a direct effect on crash frequency, as well as an indirect effect on crash frequency mediated through UFOV.

The most important issue associated with the use of LISREL models is the assessment of the fit between

FIGURE 3. LISREL model for predicting crash frequency in elderly drivers. The data from the current study were used to evaluate the model portrayed in Figure 1, but this time LISREL was used. LISREL goes beyond regression analysis in that it evaluates the independent variables in terms of whether they have direct or indirect effects on the dependent variable. (In the regression model in Figure 1, the visual function variables were grouped together into a single independent variable. In the LISREL approach depicted in Figure 3, central vision and peripheral vision are given separate status in the model; see text for further details.) The solid arrows represent the hypothesized direct effects, and each is labeled with a standardized path coefficient. Significant direct effects are indicated with an asterisk. Curvilinear lines on the left side of Figure 3 indicate that central vision, peripheral vision, and eye health are intercorrelated; the Pearson correlation coefficients label each curve (see also Table 1). UFOV and mental status were the only variables that had direct effects on crash frequency. The overall LISREL model accounted for 74% of the variance in the data, and 28% of the crash-frequency variance. Other models were considered (see text), but the model portrayed here clearly maximizes the prediction of crash frequency in our data set.
the hypothesized model and the sample data. The goodness of fit can be determined by examining the squared multiple correlation (R²), known as the coefficient of determination, for all the observed variables jointly considered. A value of 0.74 was obtained for the model in Figure 3. In addition, LISREL evaluated the goodness of fit for the model as a whole, χ² (3) = 0.44 (P = 0.732), indicating that the overall fit for the model was excellent. Finally, LISREL also provides several indices of the significance of individual parameters in the model, and a modification index that indicates whether or not model prediction can be improved by a respecification of a particular hypothesized relationship. No modification indices were significant for the hypothesized model, indicating that modifying pathways in the model would not improve it.

Although the overall model accounted for 74% of the variance in the sample data, it was also of interest to determine the R² associated with crash prediction alone. Only two variables, UFOV and mental status, directly affected crash frequency, jointly accounting for 28% of its variance. Even when we respecified the LISREL model so that central and peripheral vision were forced to have direct effects on crash frequency (in addition to their indirect effect through UFOV), there was still no increase in the amount of crash variance accounted for. The main role of central and peripheral vision in the model is their significant direct effect on the size of the UFOV; together central and peripheral vision accounted for 30% of the UFOV variance. Not surprisingly, visual attentional skills crucially depend on the integrity of information entering through the visual sensory channel. With respect to eye health, although by itself it did not significantly impact UFOV, it may have exerted an indirect effect on UFOV through its association with central and peripheral visual function (indicated by the curved lines on the left side of Figure 3).

In summary, eye health and visual function do not contribute any unique variance to crash frequency in addition to their indirect effect through UFOV. Mental status also had a significant direct effect on UFOV, and a small, but statistically significant direct effect on crash frequency as well. However, the effect of mental status on crash frequency was primarily indirect, because removal of its direct effect in the LISREL model only slightly reduced the amount of crash frequency variance accounted for (from 28% to 27%). These results thus supported our hypothesis that UFOV is a mediating variable between crash frequency on the one hand, and eye health, visual function, and mental status on the other. If the LISREL model is respecified so that UFOV is entirely removed from the model, R² only increases to 16%. Therefore the model presented in Figure 2, which includes UFOV and accounts for 28% of the crash variance, clearly maximizes the prediction of crash frequency.

One useful way of characterizing any particular predictor variable is in terms of its ability to identify drivers who have a history of crashes versus those who do not. To determine whether a particular independent variable can adequately make this discrimination, we varied our criteria of “good” and “bad” performance for each independent variable, and then using each independent variable, sorted drivers into four categories: drivers with good performance on the independent variable (low risk) with no crash history, drivers with good performance with a crash history, drivers with poor performance (high risk) with no crash history, and drivers with poor performance with a crash history. For example, to use acuity for crash prediction, a criterion of 20/40 could be adopted such that any driver with acuity of 20/40 or better would be classified as having a low crash risk, and anyone with acuity worse than 20/40 would be classified as having a high crash risk. Hits and false alarms can then be determined, where a hit represents the correct categorization of a crash-involved driver as at risk, and a false alarm represents the incorrect categorization of a crash-free driver as at risk. By varying the criterion for “poor” and “good” acuity, and determining the probabilities of hits and false alarms at each cutoff, an ROC curve (probability of hits plotted against probability of false alarms) for acuity can be generated as shown in Figure 4. Figure 4 displays ROC curves for several of the variables in our study: acuity, contrast sensitivity, peripheral vision loss, mental status, UFOV, and chronological age. Values on the diagonal indicate an equal probability of hits and false alarms, ie, an inability to classify drivers appropriately. Greater distance between an ROC curve and the diagonal corresponds to a higher sensitivity in correctly identifying drivers at risk for crashes. Figure 4 clearly indicates that the UFOV was much better at identifying crash-involved older drivers than were the other independent variables evaluated.

This study permits us to evaluate some of the popular hypotheses and stereotypes about which factors place an older driver at risk for crash involvement. As discussed earlier, visual sensory impairment in later life is often suggested as the primary cause of older adults’ higher crash rate. Visual deficits in central and peripheral vision were significantly correlated with increased crash frequency (see Table 1). For example, older drivers with severe sensitivity loss in both eyes had twice the number of crashes than did older drivers with normal visual field sensitivity, consistent with Johnson and Keltner’s earlier report. However, it is
Visual Factors and Vehicle Crashes in the Elderly

FIGURE 4. ROC curves (probability of hits plotted against probability of false alarms) for selected independent variables, which provide information about the ability of the independent variables to identify drivers who have a history of crash problems. Recall that our independent variables were performance-based tests evaluating various aspects of the visual processing system. For the purpose of generating an ROC curve for each independent variable, the definition of good performance on each variable was varied. Then, for this set of definitions, the probability of a hit was plotted against the probability of a false alarm. A hit is defined as a driver who performed poorly on the independent variable (eg, poor acuity) and had at least 1 crash on record. A false alarm is defined as a driver who performed poorly on the independent variable (eg, poor acuity), but who nevertheless had zero crashes on record. ROC curves (and their respective $d'$ values) for the following independent variables are included in Figure 4: acuity ($d' = 0.24$), contrast sensitivity ($d' = 0.67$), peripheral field sensitivity loss ($d' = 0.60$), mental status ($d' = 0.59$), and UFOV ($d' = 2.27$). We also included chronological age ($d' = 0.58$), because an age cut-off has been suggested to restrict older adults' driving privileges. It is clear that UFOV is superior to all other variables in identifying crash-involved drivers.

Important to emphasize is that no cutoff criterion in acuity, contrast sensitivity, or peripheral vision could be adopted that would place persons in the high-risk category without including a significant number of crash-free drivers in this category as well (see Figure 4).

Another common hypothesis is that older adults' crash problems are primarily due to cognitive confusion associated with dementing disease. Indeed mental status did have a significant, but small, direct effect on crash frequency in our study. However, mental status is also strongly related to performance in the UFOV task, which itself has the strongest relationship to crashes in our model. Figure 5 illustrates the relationship between UFOV and crash frequency for older drivers with good versus poor mental status. The overall crash rate is indeed slightly higher for the poor mental status group. However, the association between UFOV reduction and increased crash frequency was observed in both the good mental status group and the poor mental status group. This pattern of results is consistent with the ROC curve in Figure 4, which illustrates that mental status does not adequately identify drivers at risk for crash involvement. Furthermore, mental status had sensitivity and specificity values (0.61 and 0.62, respectively) that were markedly less than these values for UFOV (see Table 2).

Yet another popular reason cited for increased vehicle crashes in the elderly is old age itself, that is, the biologic deterioration presumed to be associated with advancing age. Although the correlation between chronological age and crash frequency was meaningless in this study due to our sampling technique, we were able to evaluate this relationship by stratifying our sample into three age groups. Figure 6 illustrates that the association between UFOV and crashes is similarly strong within each of three age groups evaluated ($F < 1.0$). Although both UFOV reduction and crashes are more prevalent with increasing age, the ROC analysis indicates that UFOV reduction is substantially better than chronological age at differentiating drivers who are at risk for crashes from those who are not.

Figure 7 illustrates that the average number of crashes increases with increasing severity of UFOV reduction. We also examined the utility of UFOV using varying cutpoint criteria. The cutpoint of 40% reduction appeared to provide the best discrimination, as
TABLE 2. Number of Subjects in High Versus Low Risk UFOV Categories Stratified by Crash History*

<table>
<thead>
<tr>
<th>UFOV Category</th>
<th>≥ 1 Crashes</th>
<th>0 Crashes</th>
</tr>
</thead>
<tbody>
<tr>
<td>UFOV Reduction &gt; 40%</td>
<td>142</td>
<td>25</td>
</tr>
<tr>
<td>UFOV Reduction ≤ 40%</td>
<td>18</td>
<td>109</td>
</tr>
<tr>
<td>Sensitivity = 89%</td>
<td>Specificity = 81%</td>
<td></td>
</tr>
</tbody>
</table>

* To evaluate the usefulness of the UFOV task as a diagnostic test for identifying crash problems, sensitivity and specificity were computed. Subjects were divided into two groups based on the percentage of UFOV reduction (<40% versus >40%; see text). Subjects in each of these groups were then divided into two categories based on crash history. Sensitivity was 89% (given a subject was crash-involved, the probability of having a UFOV reduction > 40% was 0.89). Specificity was 81% (given that a subject was not crash-involved, the probability of having a UFOV reduction ≤ 40% was 0.81).

indicated in Figure 4. Table 2 shows the frequency of subjects in each of four categories. The UFOV test had both high sensitivity (89%) and high specificity (81%) with respect to driver classification. Furthermore, the information in Table 2 was also used to calculate an odds ratio, which indicated that persons with UFOV reduction greater than 40% were six times more likely to be at least partially responsible for a crash than were those with minimal or no UFOV reduction. It should also be pointed out that of the 25 false-positive predictions, 19 were subjects who reported that they avoided driving in general, avoided driving alone, and/or avoided left turns, which thus minimized their driving exposure. In fact, if we exclude those persons who specified on the driving habits questionnaire that they avoided these particular aspects of driving, the correlation between UFOV and crash frequency increased from r = 0.52 to r = 0.62. Although avoidance is a somewhat difficult construct to measure, it appears that some older drivers are effectively compensating for visual or attentional decline and that some valid measure of exposure or avoidance behavior would be an appropriate addition to a predictive model of crash frequency.

Our sampling strategy was designed to under-represent the number of persons with zero crashes in the general population because the aim was to identify predictors of high accident frequency, and thus drivers with crash problems had to be well represented in the sample. One potential criticism of our results might be that this sampling strategy inflated our correlation coefficient between UFOV and accident frequency. Because crash frequency in the general population has a Poisson distribution, the maximum correlation of any predictor variable and crashes is significantly less than 1. Therefore, because we under-represented zero-crash drivers in the sample, our obtained correlation may be inflated relative to what it would be with a random sample. While this may be the case, it is also true that this same advantage would accrue to all the independent variables in the study and thus the overall pattern of results would not change. In fact, the correlations reported here are stronger than those reported in earlier work, probably because of the sampling technique and our ability to "clean" our dependent measure of crash frequency somewhat through elimination of those crashes that were clearly not the fault of the driver included in the
model. Furthermore, our goal in this study was to refine our ability to make predictions about individual driving performance, and not to conduct an epidemiological study on the prevalence of visual/cognitive function in the general population of older drivers. To gauge the strength of the relationship in the general population between UFOV reduction and vehicle crashes, we determined that the percentage of drivers in our reference population with no crashes in the prior 5-year period was approximately 80%. When we increased the weight of the zero crash drivers in our sample to 80%, we did indeed find that the correlation between UFOV and crash frequency decreased from 0.52 to 0.41, yet was still much stronger than in earlier studies on vision and vehicle crashes.5,7

Another question that arises is whether UFOV reduction is predictive of only certain categories of at-fault crashes (eg, failure to notice a traffic signal, merging), or whether the prediction more generally applies to many crash types. To evaluate this question, the three raters mentioned earlier were asked to classify the 364 at-fault crashes incurred by our sample into six types: failure to notice a traffic control device (n = 35), failure to notice another vehicle (n = 174), merging (n = 51), hitting the rear of another vehicle (n = 54), backing up into another vehicle or object (n = 26), and other (n = 24). The correlation between UFOV and crash frequency was similarly high for each crash type (r = 0.45 to 0.48), and the slight reduction in the strength of these correlation coefficients (compared to the analysis on all crash types) was due to decreased sample size for each type of crash, and to a lower total accident frequency for a given subject. An alternative breakdown of crashes into intersection (n = 220) and nonintersection crashes (n = 144) also revealed that the UFOV was a good predictor of both types (r = 0.41 for nonintersection, and r = 0.49 for intersection accidents). These analyses imply that the UFOV task assesses some critical visual attentional factor common to many types of at-fault crashes.

**DISCUSSION**

These results indicate that a measure of visual attention—the size of the useful field of view—is highly sensitive and specific in predicting what older drivers are at risk for crash involvement. Older drivers with a severe restriction in the spatial area over which they could rapidly use visual information were six times more likely to have incurred one or more crashes in the previous 5-year period, than were those with minimal or no restriction. There are several types of mechanisms that could potentially underlie a restriction in the size of the useful field of view in older adults. Earlier research has demonstrated that many older adults have deficits in selective attention and divided attention,17,18 as well as a slowing in the rate of visual information processing.38 These types of deficits could contribute to a narrowing of the “perceptual window.” Another potential cause of useful field of view restrictions is visual sensory impairment, such as severe loss in central and/or peripheral vision; an observer cannot attend to a visual event that is not adequately registered. Visual sensory and cognitive deficits in older adults can occur separately or together. For example, we have previously shown that useful field of view shrinkage can occur even in older adults with excellent visual field sensitivity.20 In fact 41% of our subjects in the current study with a UFOV reduction greater than 40% had an average loss of visual field sensitivity of less than 2.5 dB. Furthermore, 43% of our subjects with acuity better than 20/20 had a UFOV reduction of greater than 40%. Thus although visual status is related to the size of the useful field of view (see Table 1), good visual status alone is not a sufficient condition for a normal useful field of view. It is important to point out that because the useful field of view test relies on both visual sensory and cognitive skills, it provides a more global measure of visual functional status than either sensory or cognitive tests alone, thus improving its sensitivity and specificity in identifying older drivers at risk for crashes. We are currently examining these visual sensory and cognitive mechanisms to sort out their relative contributions to useful field of view restrictions, as well as their interactions.

Our data also imply that current visual screening techniques, such as tests of acuity and peripheral vision as used at driver licensing sites, are not adequate in identifying which elderly drivers are likely to be involved in crashes. Screening tests of acuity and peripheral vision administered at licensing offices may have other benefits for the older adult population, such as screening out those with profound vision loss and designating those in need of referral for eye care. But our analysis indicates they do not successfully identify elderly drivers who have a recent history of crash involvement, thus posing a safety risk to themselves and other road users.

Older adults in our sample covered a wide range of visual sensory abilities, as illustrated in Figure 2. For example, letter acuity among our subjects ranged from 20/10 to worse than 20/200. Our data, however, do not address the issue of whether profound acuity loss beyond 20/200 is strongly associated with crash involvement because only five subjects were in this most severe acuity-loss category. Although all five of these subjects had at least one crash in the previous 5-year period, it is interesting to note that of the older drivers in our sample with severe crash problems (3 or more crashes in prior 5-year period), only 15% had acuities worse than 20/50. Thus, at least for the older...
driver population, frequent crashes do not necessarily imply poor acuity.

Although visual attentional problems are more prevalent in the older adult population, chronological age itself did not successfully distinguish between older drivers with a history of crash problems and those who were crash free. Thus, this study clearly indicates that any policy to restrict driving privileges based solely on age is not scientifically well-founded. Decisions on the suitability of licensure in the older adult population are more appropriately based on an objective performance measure, given the diversity of functional capabilities in the elderly. A test of visual attention, such as the useful field of view, may be such a measure, given its high specificity and sensitivity for identifying drivers with a history of crash problems. The mental status test had weaker sensitivity and specificity compared to the UFOV, and furthermore, it did not significantly improve the predictions based on the UFOV test. This finding lends little support to the idea that tests of cognitive status could be widely useful in making decisions about driving licensure for the elderly, as some have suggested. Finally, the interrelationships among different aspects of the visual information processing system would not have been revealed (see Figure 3) if we had resorted to the conventional approach of studying visual sensory variables or cognitive variables in isolation.

Future work will address a number of important questions remaining. The ability of UFOV to predict future crash problems in older drivers is currently being evaluated in a prospective study using the basic design of the current study. Medical variables are also included in this study, because it has been suggested that driving difficulty in the elderly may be associated with medication usage, cardiovascular disease, and musculoskeletal/motor disorders. In addition, little is known about how visual attentional skills relate to actual driving behavior, as assessed by a road test, or to performance in a driving simulator, and how these measures could be used to better understand driving difficulties in the elderly.

It is likely that some older adults will have to be restricted from driving because of serious and irreversible deterioration in skills crucial to driving. However, it may be the case that many older adults with driving problems can improve their driving skills through treatment of ocular conditions that impair visual function (eg, cataract, glaucoma) or through training or educational programs. We have previously shown that a reduction in UFOV size in some older adults can be at least partially reversed through a training program. This training required only a modest investment in time, and the field expansion which resulted was still maintained 1 year after completion of the program. The intriguing question that remains is, given that UFOV shrinkage is associated with increased crash frequency, would expansions in UFOV size through a laboratory training program lead to improved driving performance and decreased crashes? Our study has also indicated that some older adults who are at risk for crashes because they have serious visual impairment modify their driving behavior by avoiding exposure to challenging driving situations (eg, driving alone, turning left across traffic, driving at night). This self-regulation of driving behavior was associated with a lower crash frequency. Therefore it is possible that if older adults were better educated about their visual information processing problems including visual attentional deficits, some older adults might voluntarily impose restrictions on their driving behavior that could lower their crash risk. Research to evaluate these potential interventions should be of high priority, given society’s need to enhance the mobility and personal independence of older adults, without sacrificing safety concerns.

Key Words
aging, vision, driving, crashes, visual attention

Acknowledgments
The authors thank the Alabama Department of Public Safety for providing crash data; the Eye Foundation Hospital, Birmingham, Alabama for providing facilities; Pam Alverson, Mark Graves, Mike Lewellen, Anna Harned, Dewanna Keeton, and Rebel Dahl for assistance with data collection; Milton White, Donald Turnbull, and Harold Skalika for ophthalmological exams; and Robert Sekuler, Denis Pelli, Charles Katholi, Kent Higgins, Russ Hamer, Robin Barr, Aries Arditi, and two anonymous reviewers for comments on an earlier version of this paper.

References
7. Shinar D. Driver Visual Limitations: Diagnosis and
Visual Factors and Vehicle Crashes in the Elderly


