Visual Search for Traffic Signs: The Effects of Clutter, Luminance, and Aging

Geoffrey Ho, Charles T. Scialfa, Jeff K. Caird, and Trevor Graw, University of Calgary, Calgary, Alberta, Canada

Latency and eye movement measures were used to examine the effects of aging, clutter, and luminance on visual search for traffic signs embedded in digitized images of driving scenes. Initially 14 older and 14 younger observers classified daytime and nighttime traffic scenes as containing low or high amounts of clutter. Next, an independent sample of 14 younger and 14 older participants searched for traffic signs contained within these scenes. Errors were more common among the elderly. Search efficiency declined with increased clutter and with aging. However, relative to the young, older adults did not suffer disproportionately as a result of increased clutter. The methods developed might be profitably employed to assess sign conspicuity and sign acquisition during driving.

INTRODUCTION

Safe driving behavior often depends on the ability to rapidly search visually dense scenes, extract critical information from them, and respond in an appropriate manner. Frequently this critical information is provided by signage that is intended to warn drivers of upcoming hazards or to regulate their behavior in response to anticipated traffic conditions. At times, these signs are detected and processed easily, as when an isolated sign is presented in an expected and well-illuminated location. There are other times, however, when the task is made more difficult by poor illumination, materials degradation, or visual clutter.

Visual clutter, the nontarget information in a scene, can hinder sign acquisition in a bottom-up or top-down manner (Cole & Hughes, 1984; Engel, 1971, 1976; Hughes & Cole, 1986; Shoplaugh & Whitaker, 1984). Boerema, Zwaga, and Adams (1988) suggested that longer latencies for high-clutter scenes resulted from a greater number of fixations needed to locate a target. Perhaps relatedly, Miura (1990) argued that clutter narrows the observer's useful field of view (Owsley, Ball, Sloane, Roenker, & Bruni, 1991) and increases the time needed to process the information in a given fixation.

If younger drivers have difficulty searching for traffic signs in cluttered scenes, older drivers are likely to have even greater problems (Kline et al., 1992; Schieber & Goodspeed, 1997). The elderly experience myriad visual changes that increase their susceptibility to clutter (see Kline & Scialfa, 1997). They consistently exhibit a reduced useful field of view (Owsley et al., 1991) and therefore make a larger number of eye movements to search a scene (Scialfa, Thomas, & Joffe, 1994). They have visual search deficits, particularly when target-distractor similarity is high and when there are many distractors (Plude & Doussard-Roosevelt, 1989; Scialfa & Joffe, 1997; Scialfa, Thomas, & Joffe, 1994). In addition, the elderly demonstrate cognitive declines that may affect sign acquisition, including problems with working memory and attention (Caird & Chugh, 1997; Fisk & Warr, 1998; Kidder, Park, Hertzog, & Morrell, 1997; Parasuraman & Nestor, 1991; Plude & Doussard-Roosevelt, 1989; Ponds, Brouwer, & van Wolfswinkel, 1988; Stine & Wingfield, 1990). Older adults also experience greater perceptual difficulties.
in low luminance (Chrysler, Danielson, & Kirby, 1996; Kline et al., 1992; Sivak, Olson, & Pastalan, 1981), perhaps because of reductions in retinal illuminance (Weale, 1961).

Although the aforementioned basic research is often used to predict age deficits in real-world settings, research that directly investigates aging and the search for traffic signs is limited. Recently, Schieber and Goodspeed (1997) asked observers to localize either a speed limit or business district sign that was superimposed optically on images of nighttime traffic scenes. Both the amount of clutter in the scene and the luminance of the signs were varied. Response times (RT) increased with clutter and varied inversely with luminance, and these two effects were greater in the elderly. All observers took longer to respond when a low-luminance sign was presented in a more cluttered scene, but this effect was no greater among older than among younger observers.

Even scarcer are data examining eye movements made by older adults as they search the roadway. Maltz and Shinar (1999) reported that relative to younger adults, older participants made significantly more fixations when scanning driving scenes. They also found that the distribution of fixations across a driving scene differed between older adults and younger adults, older adults had less evenly distributed fixations and more revisits than did younger adults.

In the present study, older and younger participants were asked to place daytime and simulated nighttime traffic scenes in categories corresponding to high or low visual clutter. Then, using only the scenes that were reliably categorized, we examined the effects of age, clutter, and scene luminance on sign acquisition. The foregoing discussion led to the hypothesis that the elderly would show greater clutter effects than would their younger counterparts. We also expected that nighttime scenes would be associated with greater clutter effects, which would be particularly large among the older adults.

**EXPERIMENT 1**

Previous studies have operationalized clutter either by using general descriptions of the roadway (Hughes & Cole, 1984; Schieber & Goodspeed, 1997; Shoptaugh & Whitaker, 1984) or by developing a computer-based approach based on pixel contrast (Cathcart, Doll, & Schmieder, 1989). In the present study, we operationalized clutter using subjective ratings. This approach may capture more variance among naturally occurring scenes than is possible using general descriptions of the roadway (e.g., "rural" or "urban"). Subjective ratings also allow for the top-down influences that determine clutter; for example, ignoring objects far above the horizon or vegetation on the side of the roadway.

Participants were asked to judge whether digitized images of traffic scenes were low or high in visual clutter. The results of this categorization were used to implement the factorial examination of clutter effects in Experiment 2.

**Method**

Participants. Traffic scenes were randomly presented to two groups: one group of 14 older adults ($M = 64.71$, range = 56–71 years) and one of 14 younger adults ($M = 23.43$, range = 20–27 years). Half of each age group was presented with daytime scenes, and the second half was presented with nighttime scenes. The group presented with daytime scenes was composed of 5 women (4 younger, 1 older) and 9 men (3 younger, 6 older), and the group presented with nighttime scenes was composed of 10 women (5 younger, 5 older) and 4 men (2 younger, 2 older). Younger observers were undergraduate students, and older adults were volunteers from the community. Everyone was paid $10 (Canadian) for his or her participation.

Acuity was significantly better for younger adults, as can be expected ($p < .001$). All of the younger adults had or exceeded 20/20 visual acuity, as did all but five older adults. Among these five older observers, the worst acuity was 20/30, exceeding the acuity necessary to obtain a driving license. Although all participants had been driving for at least 2 years, older adults had significantly more driving experience ($p < .001$). There were no age differences in years of education.

Stimuli. The 37 signs used for the fixation screen (i.e., to define the trial target) included regulatory signs (e.g., a stop sign) and warning signs (e.g., a yield sign). They were either created
using Microsoft Paint (Redmond, WA) or downloaded from the Internet (Moeur, 1998). Those downloaded from the Internet were converted into PCX format and resized so that they could be readily interpreted and used as stimuli for Experiment 2.

The mean length of each sign used as a fixation stimulus was 2°, and the mean width was 1.9°. Average sign luminance was 57.73 cd/m². Positioning of the signs used as fixation stimuli corresponded to the center of the roadway in the subsequent image; that is, they were placed centrally along the horizontal axis of the screen and approximately 3.81° below the horizontal meridian on the vertical axis.

Low-clutter scenes often contained only the target sign, with no more than one other sign in view. High-clutter scenes had several objects, including other signs and vehicles in proximity to the sign in both the horizontal and vertical planes. Figure 1 provides gray scale examples of low- and high-clutter images in daytime and nighttime conditions.

Traffic scenes were photographed in urban and residential areas of Calgary, Alberta, Canada. All photos were taken from the driver's seat, 1.16 m above the road, between 8:00 a.m. and 2:00 p.m. under sunny conditions. All exposures were taken at a fixed distance of 18.3 m from the target traffic sign.

Slides of the traffic scenes were digitized and given to a professional computer graphics illustrator, who manipulated the daytime scenes to simulate nighttime luminance conditions. Daytime and nighttime images were then converted to PCX format and reduced to 256 colors.

**Apparatus.** Photographs were taken in a 1991 Pontiac Tempest using a Minolta SRT 202 and Kodak Ektachrome Elite 200 film. Light readings were taken with a Minolta LS 100 photometer. Slides were digitized using a Power Mac 7100/80 computer and a Nikon 35 mm Film Scanner LS-100. Adobe Photoshop 4.0 was used to manipulate the digitized pictures, which were saved on CD-ROM using a Kodak PhotoCD writer. Images were presented on a Sony Trinitron Multiscan CPD-100 GS 15-in. monitor connected to a 486 platform. Monitor resolution was set at 640 x 480 pixels, and the refresh rate was 60 Hz.

*Figure 1.* Examples of daytime and nighttime, low-clutter (top) and high-clutter (bottom) traffic scenes used in Experiments 1 and 2. The target in the low-clutter scene is the Pedestrian Cross sign and the target in the high-clutter scene is the No Left Turn sign.
Procedure. Each observer was presented traffic scenes on the computer screen and was required to place them in either high- or low-clutter categories with respect to the target traffic signs. The test distance was maintained at 50 cm with a chin rest. Observers wore their own glasses. To avoid carryover effects, one group rated daytime scenes and another group rated nighttime scenes. For each trial the initial screen consisted of a traffic sign on a white background that provided participants the target traffic sign for which they had to judge the amount of surrounding clutter. Once they were ready, they pressed any key, and the traffic scene appeared. Participants were instructed to press the “High” key if they judged the scene to be high clutter and to press the “Low” key if they judged the scene to be low clutter. Keyboard keys were labeled as “High” (the q key) and “Low” (the p key) to avoid confusion. Each participant saw the images in a random order. A demonstration using two high-clutter scenes and two low-clutter scenes was provided before participants began.

Results

Subjective clutter was treated as a dichotomous variable (0 = low; 1 = high). Then, aggregated over the 28 observers, each image could have a score from 0 to 28. Images that had scores between 0 and 8 were categorized as low clutter, whereas images that had scores between 19 and 28 were categorized as high clutter; any images that had scores between 8 and 19 were categorized as intermediate clutter. Of the 37 images initially presented, 21 were rated as low clutter, 5 as intermediate clutter, and 11 as high clutter.

The reliability of the classifications was analyzed using Winer’s approach (1962, pp. 124–132). With daytime and nighttime images combined, the reliability of the ratings was 0.97. This result suggests that if another equally sized group judged the same images, there would be 97% agreement in the ratings of clutter. Separate reliabilities were 0.97 for the nighttime ratings and 0.91 for the daytime ratings.

Experiment 1 was carried out to determine if traffic scenes chosen to represent a bisection of the clutter dimension could be reliably cate-

gorized by groups of naive younger and older drivers. Both groups of observers consistently classified the scenes as having either low or high amounts of clutter. More important, older and younger adults were no different in their classifications. Because there were no age differences in the judged clutter, the set was used for Experiment 2.

EXPERIMENT 2

In Experiment 2, we asked whether the search for a target sign embedded in traffic scenes was systematically related to its subjective clutter and luminance and to the observer’s age. Performance was analyzed using global RT and accuracy, as well as eye movement data, which provides information not available from more traditional measures (Maltz & Shinar, 1999; Serafin, 1994; Zelinsky & Sheinberg, 1997). Highly cluttered scenes were expected to be detrimental to search performance (Shoptaugh & Whitaker, 1984), and older adults were expected to exhibit a greater clutter effect than younger adults (Plude & Doussard-Roosevelt, 1989; Schieber & Goodspeed, 1997; Scialfa & Joffe, 1997). We also hypothesized that nighttime scenes would be problematic for all observers (Sivak et al., 1981), but particularly for the elderly.

Method

Participants. Traffic scenes were randomly presented to two groups of participants. One group consisted of 14 older adults (M = 63.93, range = 54–79 years), and the other group consisted of 14 younger adults (M = 24.07, range = 18–30 years). Younger participants were volunteers from the university, and older adults responded to advertisements placed around the university campus and in community newspapers. Everyone was paid $10 (Canadian) for his or her participation.

Near acuity, near contrast sensitivity, and intraocular pressure were assessed for each participant. Younger adults had significantly better visual acuity than did older adults (p < .05), but the mean acuity for older adults was good. The poorest acuity for an older adult was 20/24, much better than the acuity required to obtain and keep a driver’s license. With respect
to contrast sensitivity and intraocular pressure, all participants were normal relative to their age. There were no age differences in education, $p > .15$, although, as expected, older adults had more driving experience, $p < .001$.

**Stimuli.** The judgment data of Experiment 1 determined the images that were chosen for Experiment 2. The 32 target signs in the low- and high-clutter scenes were compared on eccentricity, size, and luminance. These tests revealed no significant differences in any features except vertical sign size. Signs in the high-clutter category were significantly smaller vertically than signs in the low-clutter category, $t(50) = 2.16, p = .058$.

Because size is an important factor in the conspicuity of traffic signs (Cole & Hughes, 1984), several images were discarded. As a result, eight low-clutter images and three high-clutter images were eliminated for Experiment 2 (see Table 1 for a list and characterization of the signs used). The following analyses refer to the 21 low- and high-clutter images that were retained. Four images with intermediate clutter were presented to participants for exploratory purposes but did not produce systematic effects and so will not be discussed further. In total, then, there were 15 low-clutter and 8 high-clutter scenes. Each image had its nighttime counterpart, for a total of 16 high- and 26 low-clutter scenes.

**Apparatus.** All participants were fitted with a set of Burton Trial Lenses, with their own near visual correction. If no correction was needed, neutral lenses were used. Contrast sensitivity was tested using the Vistech Contrast Sensitivity Chart, and intraocular pressure was tested using a Reichert NCT2 noncontact tonometer.

Stimulus presentation (viewing distance = 50 cm) and data collection were instrumented using the same system described in Experiment 1. Eye movements were recorded using the Eyegaze Development System (EDS) and software from LC Technologies, Inc. (Fairfax, Virginia) Low-level infrared light (880 nm)

### TABLE 1: Descriptive Statistics for Signs Used in Experiment 2 as Presented on Monitor

<table>
<thead>
<tr>
<th>Sign</th>
<th>Height (°)</th>
<th>Width (°)</th>
<th>Screen Luminance (cd/m²)</th>
<th>Eccentricity (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Low Clutter</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80 km Ahead</td>
<td>2.04</td>
<td>1.53</td>
<td>17.89</td>
<td>8.22</td>
</tr>
<tr>
<td>Added Lane</td>
<td>1.91</td>
<td>1.91</td>
<td>27.04</td>
<td>5.71</td>
</tr>
<tr>
<td>Do Not Enter</td>
<td>1.65</td>
<td>1.40</td>
<td>26.01</td>
<td>4.57</td>
</tr>
<tr>
<td>Keep Right</td>
<td>0.89</td>
<td>0.76</td>
<td>26.83</td>
<td>5.96</td>
</tr>
<tr>
<td>Median</td>
<td>1.65</td>
<td>0.76</td>
<td>23.74</td>
<td>5.46</td>
</tr>
<tr>
<td>Pedestrian Cross 1</td>
<td>1.91</td>
<td>1.91</td>
<td>30.38</td>
<td>6.21</td>
</tr>
<tr>
<td>Pedestrian Cross 2</td>
<td>1.27</td>
<td>0.89</td>
<td>34.36</td>
<td>5.21</td>
</tr>
<tr>
<td>Playground</td>
<td>1.27</td>
<td>1.27</td>
<td>33.00</td>
<td>4.19</td>
</tr>
<tr>
<td>Right Lane</td>
<td>1.78</td>
<td>1.15</td>
<td>23.37</td>
<td>7.84</td>
</tr>
<tr>
<td>Right Turn</td>
<td>1.02</td>
<td>1.15</td>
<td>22.28</td>
<td>4.95</td>
</tr>
<tr>
<td>Stop</td>
<td>0.89</td>
<td>0.89</td>
<td>15.25</td>
<td>6.59</td>
</tr>
<tr>
<td>Traffic Lights</td>
<td>1.91</td>
<td>1.78</td>
<td>28.79</td>
<td>6.97</td>
</tr>
<tr>
<td>Yield</td>
<td>1.02</td>
<td>1.15</td>
<td>21.14</td>
<td>7.84</td>
</tr>
<tr>
<td><strong>High Clutter</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Keep Right</td>
<td>0.89</td>
<td>0.89</td>
<td>48.60</td>
<td>5.71</td>
</tr>
<tr>
<td>No Left</td>
<td>0.89</td>
<td>1.15</td>
<td>47.05</td>
<td>6.84</td>
</tr>
<tr>
<td>Pedestrian Cross</td>
<td>1.40</td>
<td>0.89</td>
<td>39.79</td>
<td>7.22</td>
</tr>
<tr>
<td>Playground</td>
<td>0.89</td>
<td>0.89</td>
<td>13.67</td>
<td>5.08</td>
</tr>
<tr>
<td>Right Turn</td>
<td>1.15</td>
<td>0.89</td>
<td>17.48</td>
<td>7.22</td>
</tr>
<tr>
<td>Stop</td>
<td>0.89</td>
<td>0.89</td>
<td>15.45</td>
<td>7.47</td>
</tr>
<tr>
<td>Traffic Lights</td>
<td>2.16</td>
<td>2.16</td>
<td>37.49</td>
<td>6.84</td>
</tr>
<tr>
<td>Yield</td>
<td>1.02</td>
<td>1.15</td>
<td>47.19</td>
<td>5.71</td>
</tr>
</tbody>
</table>
was directed to the participant’s eye, and its reflection was measured with a Sanyo infrared-sensitive camera using a monocular, pupil-center-corneal reflection technique to track the eye at 30.3 Hz. A fixation was recorded if there were two consecutive samples within a spatial window of 11 pixels, which is approximately 0.50°. For more details on the EDS, see Scialfa, Thomas, and Joffe (1994).

Procedure. In the visual search task, participants were instructed that the first screen to appear would be a fixation screen consisting of a traffic sign. This traffic sign served as the fixation point and also indicated the sign for which to search in the subsequent image. Once they were ready, participants pressed any key, and the traffic scene image appeared. During half the trials, the traffic scene contained the sign presented in the fixation screen; during the other half, it did not. The participant’s task was to respond “present” or “absent” by pressing predetermined keys as quickly as possible. The traffic scene image remained on the screen until the participant responded. After each response, the participant was given accuracy feedback. A plus sign indicated correct, and a minus sign indicated incorrect.

Participants were initially guided through 10 practice trials that contained both target-present and target-absent scenes not used in the experiment proper. Among the 10 scenes, 5 were daytime scenes and 5 were nighttime scenes. To ensure that all participants understood the task, an accuracy of 90% had to be obtained on the practice trials before continuing to the actual testing. If the participant failed to obtain 90% accuracy, the instructions were repeated, and the practice trials were presented again.

There were two blocks of 25 trials, each preceded by calibration of the EDS, with an optional rest between them. Half the participants searched the 25 daytime scenes first and the 25 nighttime scenes second. For the remaining participants, the order was reversed. Clutter and target presence were randomly ordered for each participant. On average the procedure did not exceed 1 hr.

Results

A trial was omitted if a saccade exceeded the screen boundaries and if the participant was not properly fixating on the target sign prior to the onset of the scene to be searched. The analysis of five dependent measures is reported: (a) errors, (b) reaction time, (c) fixation number, (d) average fixation duration, and last fixation duration. For each measure the effects of age, clutter, luminance, and target presence were evaluated in an Age (2) × Clutter (2) × Luminance (2) × Target Presence (2) univariate mixed-model analysis of variance (ANOVA) with a Geisser-Greenhouse correction for violations of sphericity (Maxwell & Delaney, 1990). With the exception of the error data, all analyses omitted trials ending in an incorrect response.

Errors. On average, older adults were less accurate than younger ones, F(1, 26) = 5.99, p = .021. As shown in Figure 2, errors were more common in high-clutter scenes than in low-clutter scenes, F(1, 26) = 52.48, p < .001 and on target-present trials compared with target-absent trials, F(1, 26) = 6.16, p = .020. There was a Clutter × Presence effect, F(1, 26) = 6.77, p = .015, because more errors were made on target-present trials involving greater clutter.

There was a Luminance × Presence interaction, F(1, 26) = 5.67, p = .025, because accuracy for daytime scenes was independent of target presence, but errors for nighttime scenes were more common on target-present trials than on target-absent trials. There was also a Clutter × Luminance interaction, F(1, 26) = 10.95, p = .003, because in daytime scenes, errors were generally more common in high clutter than in low clutter, while in nighttime scenes, there was not a consistent clutter effect. Finally, there was a Clutter × Presence × Luminance interaction, F(1, 26) = 10.79, p = .003. Errors were relatively high in three of the four high-clutter conditions, but nighttime target-absent trials produced the lowest error rates of any condition tested. More important, there were no interaction effects involving age.

Reaction times. Reaction time data, shown in Figure 3, indicate that younger adults generally responded more quickly than older adults. Reaction times were also dependent on the amount of clutter and the presence of the target. The main effects of age, F(1, 23) = 27.97, p < .001; clutter, F(1, 23) = 570.07, p < .001;
and presence, $F(1, 23) = 49.49, p < .001$, were all significant, whereas the main effect of luminance was not, $F(1, 23) = 0.01, p = .912$.

In addition, several higher-order effects were significant. Clutter interacted with luminance, $F(1, 23) = 4.91, p = .037$; and presence, $F(1, 23) = 9.91, p = .005$. The Age $\times$ Presence interaction was also significant, $F(1, 23) = 7.80, p = .01$, because age differences were greater on target-absent trials than on target-present trials. The three-way interaction of age, clutter, and luminance was not significant, $F(1, 23) = 2.85, p = .105$. However, the Clutter $\times$ Luminance $\times$ Presence interaction was significant, suggesting that the Clutter $\times$ Luminance interaction was greater on target-absent trials. Older adults had even slower RTs in the target-absent condition than did their younger counterparts, as shown by the significant four-way interaction, $F(1, 23) = 10.66, p = .003$.

**Fixation number.** Fixation number is a measure that is strongly correlated with RT (Scialfa, Thomas, & Joffe, 1994) and therefore can be used as a measure of search efficiency. Fixation number was calculated by omitting the first fixation, which was experimentally determined by the fixation screen. As shown in Figure 4, the fixation number data have a similar pattern to that found in RTs. Older adults made more fixations than did younger adults, $F(1, 23) = 31.89, p < .001$. Moreover, more fixations were needed for high-clutter than for low-clutter, $F(1, 23) = 60.40, p < .001$, and for target-absent scenes than for target-present.
Figure 3. Reaction time for target-present (TP; top) and target-absent trials (TA; bottom), and younger (Y; left) and older (O; right) participants as a function of clutter and luminance.

scenarios, $F(1, 23) = 74.40, p < .001$. Once again, the main effect of luminance was not significant.

Similar to reaction time, the Age $\times$ Presence interaction was significant, $F(1, 23) = 11.53, p = .002$, as was the interaction of Clutter $\times$ Presence, $F(1, 23) = 12.76, p = .002$. Nighttime scenes with high clutter also required more fixations than did daytime scenes, $F(1, 23) = 9.39, p = .005$. No other higher-order interactions reached significance, although the Clutter $\times$ Luminance $\times$ Presence interaction was marginally so, $F(1, 23) = 3.19, p = .087$.

Last fixation duration. The last fixation duration reflects the comparison of the last fixated object with the target representation and the terminal decision (in our case, a keypress) regarding target presence. Age differences in this last stage of processing might accrue because of difficulties with the comparison process itself (Scialfa & Thomas, 1994) or because the elderly are more cautious in making overt responses (see Salthouse, 1982).

The data for last fixation duration are shown in Figure 5. The main effect of age was significant, $F(1, 23) = 15.87, p = .001$, as was the main effect of presence, $F(1, 23) = 23.41, p < .001$. The main effect of luminance was not
Figure 4. Fixation number for target-present (TP; top) and target-absent (TA; bottom) trials, and younger (Y; left) and older (O; right) participants as a function of clutter and luminance.

significant, $p = .127$, although the main effect of clutter was marginal, $F(1, 23) = 3.12, p = .091$. The Age x Presence interaction was significant, $F(1, 23) = 8.71, p = .007$, but, in contrast with the RT data, age differences were greater on target-present trials than on target-absent trials. There was a marginal effect of Clutter x Luminance x Presence, $F(1, 23) = 3.24, p = .085$, suggesting that the Clutter x Luminance interaction was greater on target-present trials. No other effects were significant, $p > .12$.

Average fixation duration. Average fixation duration, shown in Figure 6, is a measure of the amount of time needed for information processing at each fixation. Average fixation duration did not include the first fixation duration. Previous studies have found that the first fixation to initiate a saccade may be quite different from subsequent fixations, as it is much longer and involves time to process the whole display and initiate a first saccade (Scialfa & Joffe, 1998; Zelinsky & Sheinberg, 1997).

Younger adults had shorter fixation durations than did older adults, and low-clutter scenes resulted in shorter fixations than did high-clutter scenes. There was a main effect of age, $F(1, 23) = 11.11, p = .005$; clutter, $F(1, 23) = 130.08, p = .001$; and presence, $F(1, 23) = 49.43, p < .001$, each in the anticipated
direction. Higher-order analyses revealed a significant Age x Clutter interaction, $F(1, 23) = 5.76$, $p = .025$. This interaction occurred because younger participants exhibited longer fixations in high-clutter scenes than in low-clutter scenes, whereas older participants exhibited fixation durations that were independent of clutter. A Clutter x Presence x Luminance interaction, $F(1, 23) = 9.85$, $p = .005$, was obtained as well. Relative to daytime scenes, high-clutter nighttime scenes containing a target resulted in longer average fixation durations. No other effects were significant, $p > .12$.

**Discussion**

This study was intended to determine if observers could reliably estimate visual clutter in driving scenes and if these estimates of clutter would predict the efficiency of visual search for traffic signs by younger and older drivers. In Experiment 1 we found that observers were able to classify scenes reliably on the basis of clutter. In Experiment 2 these classifications allowed for accurate predictions of search performance. Those scenes classified as high in clutter required longer latencies and more fixations to acquire the sign, were associated with more errors, and had longer fixation durations than did those that were low in clutter. Luminance effects were negligible, although high-clutter nighttime scenes did impair search efficiency on RT and fixation number measures. In addition, search was impaired on target-absent trials for both age groups.
In Experiment 2, older adults were less accurate and slower than younger adults and executed more eye movements to acquire signs. The age effects on RT and fixation number were more pronounced on target-absent trials. In contrast, the last-fixation data exhibited an Age × Presence interaction in the opposite direction. That is, older adults used the visual cues that discriminated targets and distractors to quickly isolate the target on target-present trials. However, they took longer to correctly decide that a target sign was not present on target-absent trials, and they also took longer to make a terminal decision regarding the sign's match to the trial target on target-present trials. By devoting more time to searching for something that is not present, drivers will necessarily limit the time devoted to performing other tasks, and this effect may be more pronounced for older adults (Ponds et al., 1988).

Interestingly, older adults in the present study were not more adversely affected by increased clutter than were the young. In contrast, Schieber and Goodspeed (1997) reported age deficits in sign acquisition in more cluttered driving scenes. An important difference between this study and that of Schieber and Goodspeed rests in the operational definition of clutter. They defined a low-clutter scene as an "isolated 2-lane highway,” a moderately cluttered scene as “a typical street in a commercial district of a small city,” and a high-clutter scene as a “downtown street of an urban area complete with scores of commercial establishments with illuminated business signs” (p. 1365). In the present study, all scenes classified as low clutter were either residential or quiet commercial roads in a city, whereas those classified as high clutter were urban roadways with several competing signs and other visual

Figure 6. Average fixation duration for target-present (TP; top) and target-absent trials (TA; bottom), and younger (Y; left) and older (O; right) participants as a function of clutter and luminance.
distractors. Thus low-clutter images in the present study were most similar to moderately cluttered images in the Schieber and Goodspeed study, whereas the two studies’ high-clutter conditions were roughly comparable. Examination of Schieber and Goodspeed’s Figure 4 (p. 1365) suggests that there is no Age × Clutter interaction in moderate- and high-clutter conditions. Thus roughly equivalent clutter conditions in these two studies show the same lack of an Age × Clutter interaction.

This study also examined age differences in searching daytime and nighttime scenes. There was no main effect of lumiance, nor was there an interaction between age and lumiance. To explain this pattern in the data, consider our manipulation of lumiance. We simulated the effect of using headlamps to illuminate retro-reflective signs by decreasing overall lumiance of the scene but increasing a sign’s lumiance. Thus the lumiance contrast of the sign increased. Schieber and Goodspeed (1997) found similar results by holding background lumiance constant and varying sign brightness, creating “bright” and “ultrabright” signs. Both studies demonstrate the benefits of increasing lumiance contrast.

We are not suggesting that the search for signs is equally easy in nighttime and daytime scenes. First, we did find that highly cluttered nighttime scenes were associated with longer RTs and a greater number of fixations. Second, our manipulation optimized lumiance contrast in the nighttime condition and did not include the detrimental effects of nighttime driving, such as increased glare and a reduction in overall lumiance, both of which can affect older drivers in particular. Thus the small lumiance effects we found may not generalize to different nighttime conditions. Perhaps questions regarding nighttime conspicuity might best be left for controlled field studies (e.g., Chrysler et al., 1996).

**Recommendations for research and driving environments.** The scenes judged to be higher in clutter (Experiment 1) were associated with poorer search performance (Experiment 2). These findings indicate that the subjective clutter ratings provide valid measures of visual clutter in naturalistic scenes. Nominal estimations of clutter based on scene type (e.g., Schieber & Goodspeed, 1997) and computational approaches (e.g., Cathcart et al., 1989) do not take into account experience, expectations, and other top-down factors that can influence scene perception generally (Henderson, Weeks, & Hollingsworth, 1999) and sign conspicuity in particular. Thus the successful use of subjective clutter ratings in the present study provides an alternative method of evaluating clutter in applied research.

However, obtaining a representative set of subjective clutter ratings can be expensive and time-consuming. As such, further research might examine the relation between subjective and computational clutter ratings, along with their combined ability to predict sign acquisition. If subjective ratings of clutter do not add to the variance in performance accounted for by computational measures alone, then they might not be needed in many applications. On the other hand, if incremental variance in sign acquisition is accounted for by subjective clutter measures, then it might be profitable to explore perceived clutter measures in detail in order to enhance signage and to build better computational models of sign acquisition.

Older adults were slower and less accurate and required more fixations to acquire a traffic sign. This suggests that in time-limited situations that involve visually complex scenes (e.g., a busy intersection), older adults are more likely than younger adults to misidentify a sign or miss a sign altogether. The driving population is aging rapidly. In light of this trend and the behavioral consequences reported here, roadway engineers should consider reducing the number of competing signs (e.g., advertisements), avoiding redundant signs, and making traffic signs that are central to the safety of the driver more conspicuous.

Although older adults did not show a more pronounced clutter effect, it is likely that our results present an optimistic picture of older adult’s search capabilities. Our stimuli were static scenes with high visibility. It is likely that in a more realistic driving simulation, with dynamic stimuli and varying degrees of visibility, older adults would be more dramatically affected by clutter. Also, our task required only a single keypress response and thus was not as cognitively demanding as driving. Still, although
the present study is not strongly representative of real driving, it paints a clearer picture of older adult’s search behavior of naturalistic scenes.

The extension of these studies to dynamic stimuli and field research is a logical next step. As time, space, and vehicle control constraints interact with the task of acquiring sign information, it is expected that clutter effects will become more pronounced. For example, the temporal window in which a sign can be viewed while in transit is limited. Search within a collection of signs will terminate as the window closes, and driving errors may result because critical information has not been obtained. This effect may be more problematic for the elderly, who will continue to have a growing presence on the roadway.

ACKNOWLEDGMENTS

The authors thank David Stewart for the programming that made this project possible and Scott Park for help with simulating nighttime scenes. The authors also thank Bob Dewar for his helpful comments on earlier drafts of the manuscript. These data were presented, in part, at the annual meeting of the Human Factors and Ergonomics Society, Houston, TX, in September 1999. Funding was provided in part by the Natural Science and Engineering Research Council of Canada (NSERC).

REFERENCES


Geoffrey Ho is a Ph.D. student at University of Calgary, Calgary, Alberta, Canada. He received his M.S. in psychology from University of Calgary in 1999.

Charles T. Scialfa is a professor of psychology at University of Calgary. He received his Ph.D. in psychology from University of Notre Dame in 1987.

Jeff K. Caird is an associate professor of psychology at University of Calgary. He received his Ph.D. in human factors from University of Minnesota in 1994.

Trevor Graw is a network security systems engineer at CDG, Inc., Calgary, Alberta, Canada. He received his B.S. from University of Calgary in 1999.

*Date received: November 10, 1999
Date accepted: November 7, 2000*