Experimental Evaluation of Retroreflective Markings on Trains at Grade Crossings

by

Joseph Conti

Submitted to the Department of Mechanical Engineering on May 8, 1998 in Partial Fulfillment of the requirements for the Degree of Master of Science in Mechanical Engineering

Abstract

Every year in the United States, hundreds of accidents occur at grade crossings due to motor vehicles colliding with trains. Furthermore, a large majority of these accidents take place at night in rural areas. One proposed solution to prevent such accidents involves mounting retroreflective material on the sides of trains so that a vehicle’s headlights will illuminate the reflectors and make the train more conspicuous. The objective of this research was to determine which train mounted reflector pattern gives an approaching driver the best train recognition. Four reflector patterns based on previous research were selected for this study, and a computer based nighttime driving simulator was developed for this research. In the first experiment, over a thousand scenes containing the view of a road intersection and a grade crossing were displayed to the subject. The subject’s recognition of different reflector patterns was recorded and subsequently analyzed using the Receiver Operating Characteristic (ROC) based on Signal Detection Theory (SDT). The second experiment involved a driving task in which the subject encountered numerous grade crossings, and the recognition distance between the train and subject’s position was recorded and analyzed.

Thesis Supervisor: Dr. Thomas B. Sheridan
Ford Professor of Engineering and Applied Psychology
Acknowledgments

Foremost, I would like to thank my advisor, Dr. Thomas Sheridan, for his wisdom, guidance, and trust.

Many thanks to my sponsor, Dr. Jordan Multer of the Volpe Center, for his expertise in experimental design and patience in viewing countless simulator demos.

My gratitude to Debra Williams and Anya Carrol of Volpe for their valuable input, as well as Amy Holton and Robert DiSario for their help in compiling and analyzing experimental data.

Thanks to my undergraduate advisor at the University of Maryland, Dr. Guangming Zhang, who encouraged me to pursue research and graduate school.

Entering and leaving MIT can be a very stressful process if not for the Mechanical Engineering Department’s Leslie Regan, who always has time to ease the various administrative crises graduate students experience.

Thanks to Human-Machine System Laboratory members: Santanau Basu, Jay Einhorn, Steve Landry, Helias Marinakos, Nicholas Malsch, Shinsuk Park, Jim Thompson, and Steven Villareal.

I never would have made it this far if not for Esook Yoon, a priceless friend whose love, advice, and friendship always fills me with me confidence and happiness.

I am grateful to my parents for encouraging me to pursue any dream that I set my sights on.

To any future graduate student who’s eyes fall upon this, carpe diem. Time will pass quicker than you think!
Table of Contents

ABSTRACT........................................................................................................................................2

ACKNOWLEDGMENTS .......................................................................................................................3

TABLE OF CONTENTS .......................................................................................................................4

LIST OF FIGURES ..............................................................................................................................6

LIST OF TABLES ................................................................................................................................7

1. INTRODUCTION .............................................................................................................................8
  1.1 RETROREFLECTORS .....................................................................................................................8
  1.2 PRIOR RESEARCH .........................................................................................................................9

2. EXPERIMENT I ................................................................................................................................12
  2.1 OBJECTIVE ................................................................................................................................12
  2.2 METHOD .....................................................................................................................................12
    A. Apparatus .................................................................................................................................12
    B. Signal Detection Theory ..........................................................................................................15
    C. Experimental Design ................................................................................................................16
      Independent Variable ..................................................................................................................16
      Dependent Variable ...................................................................................................................18
      Summary of Experimental Design ..............................................................................................19
    D. Subjects ....................................................................................................................................20
    E. Instructions and Treatment of Subjects ....................................................................................20
  2.3 RESULTS ....................................................................................................................................21
  2.4 DISCUSSION ................................................................................................................................23

3. EXPERIMENT II .............................................................................................................................24
  3.1 OBJECTIVE ................................................................................................................................24
  3.2 METHOD .....................................................................................................................................24
    A. Apparatus .................................................................................................................................24
    B. Experimental Design ................................................................................................................26
      Independent Variable ..................................................................................................................27
      Dependent Variable ...................................................................................................................37
      Errors .........................................................................................................................................37
    C. Subjects ....................................................................................................................................27
    D. Instructions and Treatment of Subjects ....................................................................................28
  3.3 RESULTS ....................................................................................................................................28
    A. Objective Results .......................................................................................................................28
    B. Subjective Results ....................................................................................................................29
  3.4 DISCUSSION ................................................................................................................................31

4. CONCLUSION ................................................................................................................................31
List of Figures

Figure 1. Hopper Car Reflector Patterns ................................................................. 11
Figure 2. Flat Car Reflector Patterns ..................................................................... 11
Figure 3. Truck Reflector Patterns ......................................................................... 11
Figure 4. Experiment I Apparatus .......................................................................... 13
Figure 5. Experiment I Layout .............................................................................. 14
Figure 6. View of Urban Scenery ......................................................................... 14
Figure 7. Stimulus-Response Matrix ....................................................................... 16
Figure 8. Average ROC Curves ............................................................................. 22
Figure 9. View of Driving Simulator ........................................................................ 25
Figure 10. Experiment II Apparatus ....................................................................... 26
Figure 11. Recognition Distance versus Reflector Type ......................................... 29
Figure 12. Subjective Rating versus Reflector Type ................................................ 29
Figure 13. Picture of the Steering Wheel/Pedal Assembly ....................................... 50
Figure 14. Driving Interface Connectivity ............................................................. 51
Figure 15. Overview of Software .......................................................................... 53
Figure 16. Directory Layout of the Car Simulator Software ...................................... 55
List of Tables

Table 1. Experiment I: Independent Variables................................................................. 17
Table 2. Experiment I: Dependent Variables.................................................................. 18
Table 3. Experiment I: Organization.............................................................................. 19
Table 4. Experiment II: Independent Variables............................................................... 27
1. Introduction

Defined as the intersection of a roadway and railroad tracks, grade crossings exist at over 168,000 locations throughout the United States (FRA, 1994). Unfortunately, many accidents occur at grade crossings involving the collision of motor vehicles and train consists, where a *consist* is a linkage of train cars that typically make up a complete train. In 1994 there were over 1,183 motor vehicle run into train (RIT) accidents, 613 of which occurred during nighttime conditions (U.S. Department of Transportation [DOT], 1998). These accidents usually occur in rural areas where no ambient light exists except for the motorist’s headlights. In addition, the dark color of typical freight trains and the accumulation of dust and grime makes the task of train recognition even more difficult.

The DOT has sponsored studies exploring many methods to prevent such accidents including train mounted incandescent/strobe lights, paint schemes, and audible warning devices (Carrol, Multer, & Markos, 1995; Aurelius & Korobow, 1971). But, a promising alternative method involves mounting retroreflective material on the sides of train cars so that the motorist’s headlights will illuminate the reflectors giving the driver adequate knowledge that a train is passing through the grade crossing.

1.1 Retroreflectors

Retroreflectors are a special type of reflectors that reflect light back to the source, regardless of angle. In other words, the reflected ray of light is parallel to the incident ray of light, and this allows drivers to see the reflectors from different angles. A model of retroreflectors is described by the following equation (Poage, Pomfret, & Hopkins, 1982; McGinnis, 1979).

\[ E = \frac{I_s A R t^2 T}{d^4} \]

where

- \( E \) = Illuminance received in units of lx or ft-candles
- \( I_s \) = Intensity of light toward retroreflector from source in units of candela (cd)
- \( A \) = Area of retroreflector in units of m² or ft²
- \( R \) = Specific intensity per unit area, SIA, in units of cd/lx/m² or cd/ft-candle/ft²
\[ t = \text{transmissitivity of atmosphere per distance} \]
\[ d = \text{distance from source to reflector in units of m or ft} \]
\[ T = \text{transmissitivity of windshield and headlights} \]

The variable \( R \), specific intensity per unit area (SIA), is used to compare performance of reflector material with one another; a higher reflector SIA value means greater illuminance.

**1.2 Prior Research**

The notion of increasing a train’s conspicuity through retroreflectors is not a new concept and was studied as early as the middle twentieth century (Stalder & Lauer, 1954). At that time, however, it was not feasible due to the high cost and poor performance of the available retroreflective material. Roughly twenty-five years later, the DOT sponsored more train mounted reflector studies using the then available retroreflective material: *enclosed lense* and *encapsulated lense*. Poage et al. determined that the minimum retroreflective intensity for train mounted reflectors should be about 45 cd/ft-candela in order to give an approaching driver adequate stopping distance (1982). This study also concluded that the reflective intensity of the reflector material degrades to 10% of its original value after two years of service on a train car, and although frequent washings improved the reflectors efficiency, they still eventually required replacement. Thus, the use of train mounted retroreflectors was still unfeasible.

A decade later, however, the availability of a new type of retroreflective material with a higher SIA and resistance to wear prompted the DOT to further explore the use of train mounted retroreflectors. This latest retroreflector material is called *prismatic* or *cube corner* due to the use of small prisms embedded in the material. In a study for researching the use of retroreflectors on trucks, performance of this new material did not degrade significantly over time, and regular washing of the reflectors sustained their effectiveness (Olson, Campbell, Massie, Battle, Traube, Aoki, Sato, & Pettis, 1992). In 1996, the DOT sponsored a study conducted by the University of Tennessee to explore various reflector patterns (colors and configuration) to improve the nighttime conspicuity of trains. (Ford, Richards, & Hungerford, 1996). This study concluded that:
A standardized retroreflector pattern is beneficial to train recognition.

The pattern should be made of red and white reflectors.

The pattern should not be confused with roadway signs or reflectors from other objects (i.e. trucks).

The pattern should communicate the size of the train car through outlining or an even distribution.

With that in mind the DOT’s John A. Volpe National Transportation Systems Center (Volpe Center) located in Cambridge, Massachusetts proposed four reflector patterns based on prismatic retroreflector material to conduct further human factor research using a driving simulator. These four reflector patterns: massed outline, vertical bars, variable vertical bars, and horizontal bars are presented in Figures 1 and 2 on the hopper train cars and flat train cars, respectively. These patterns were chosen to represent the spectrum of reflector layout strategies; namely distributing the reflectors along the train car (horizontal bars and vertical bars), lumping them towards the ends (massed outline), or a combination of the two (variable vertical bars). In this study all four hopper car reflector patterns each used exactly 144 square inches of red reflector material and 144 square inches of white reflector material (see Figure 1). The flat train car patterns each used exactly 72 square inches of red reflector material and 72 square inches of white reflector material (see Figure 2). The patterns in Figures 1 & 2 are all comprised of reflector strips 4 inches wide with varying length.

Therefore, the objective of this research was to determine which of the four reflector patterns in both Figures 1 and 2 best facilitates the task of train recognition using a driving simulator. In order to make recognition more difficult and realistic, both experiments included a tractor trailer truck with reflectors. Currently, the U.S. Code of Federal Regulations requires that “not less than half of the length of the [truck] trailer is covered [with reflectors] and the spaces are distributed as evenly as practicable” (49 USC 571.108). Figure 3 depicts the four truck reflector patterns utilized in this study. These truck reflector patterns were chosen to represent the typical patterns used by trucking companies and manufacturers.
Figure 1. Hopper Car Reflector Patterns

Figure 2. Flat Car Reflector Patterns

Figure 3. Truck Reflector Patterns
2. Experiment I

2.1 Objective

The purpose of the first experiment was to determine which of the four hopper/flat train reflector patterns described in Figures 1 & 2 gives the best train recognition and in general to evaluate the driver’s ability to detect and distinguish between trains and trucks.

2.2 Method

A. Apparatus

This experiment took place at the Volpe Center. The equipment used for the experiment was a Silicon Graphics Indigo2 workstation, a Barco 808S projector, and a 6x8 ft screen 15 feet away from the subject (see Figure 4). A description of the hardware and software used in this experiment can be found in Appendix C & D.
Each subject viewed 1200 computer generated scenes projected onto the screen in front. These scenes simulated the view of a driver in a stationary vehicle 500 feet away from a grade crossing and road intersection as in Figure 5. In some of these scenes, a truck or train passed through the intersection and the subjects were asked to identify the object. In the urban scenery, noise was presented in the form of buildings with lighted windows and street lights. Figure 6 provides a snapshot of the urban scenery.
Figure 5. Experiment I Layout

Figure 6. View of Urban Scenery
Each scene lasted 500 milliseconds and was followed by two questions asking what the subject saw and how confident they were of their decision.

B. Signal Detection Theory

Data were analyzed using signal detection theory (SDT). Originally developed in the early 1950’s for researching the acquisition of radar signals, SDT was first applied to the field of psychology in the late 1950’s (Green and Swets, 1988). Since then many references have incorporated the topic of SDT along with one of its most useful components, the receiver operating characteristic (ROC) into experimental psychology (Sheridan and Ferrell, 1974).

When applied to experimental psychology the basic problem of SDT is the ability of a human to discriminate a signal from noise. Given an input of either a signal (plus noise) or noise (alone), which will the subject observe? If nothing (noise) is perceived when a signal is actually present then a miss has occurred; the subject missed the signal. However, if the subject correctly observes a signal when a signal is actually present then a hit has occurred. Furthermore, if a signal is perceived when only noise is present then a false alarm has occurred. Finally, if nothing (noise) is observed when only noise is present then a correct rejection has occurred.

The matrix in Figure 7 plots these four possible outcomes and subsequently gives us the classic signal detection theory stimulus-response matrix also known as the decision matrix (Sheridan and Ferrell, 1974).
Through analyzing experimental data, the probability of hits and false alarms can be computed, and plotting these two probabilities versus each other gives us the Receiver Operating Characteristic (ROC) curve. A more in-depth description in constructing ROC curves can be found in many references (Egan, 1975; Green and Swets, 1988; Sheridan and Ferrell, 1974). The ROC curves in this experiment were based on a rating procedure which allows a single curve to be generated from only one subject, unlike a binary-decision procedure which can generate only one point on the ROC curve. Therefore, this rating procedure is more efficient than the binary-decision procedure in creating the ROC curve and arguably yields an equivalent ROC curve (Green and Swets, 1988).

C. Experimental Design

**Independent Variable**
The 1200 scenes displayed in each experiment were made up of 32 unique scenes shown repetitively and randomly throughout the experiment. Table 1 describes these 32 scenes by the variables S or N which correspond to a signal or noise, respectively. The scene was a signal, S, if and only if a train was present at the grade crossing, otherwise the scene was noise, N.
Table 1. Experiment I: Independent Variables

The first subscript of these variables, denoted by an $r$ or $u$, represents the setting of the scenery which is either rural or urban, respectively. The rural setting is simply a road with an intersection 500 ft away, whereas the urban setting is just like the rural setting with the addition of street lamps and buildings with lighted windows (see Figure 6).

The second subscript, denoted by the subscripts $h$, $f$, or $t$, represents the type of object displayed in the intersection which is either a hopper train consist, a flat train consist, or a truck, respectively. In all cases the object in the intersection is moving right to left at a constant speed of 30 mph. The third subscript, the integer 1, 2, 3, 4 or 5, represents the reflector type: massed outline, vertical bars, variable vertical bars, horizontal bars, or no reflectors at all. These reflector types are described in Figures 1, 2, and 3, and the right three columns of Table 1 summarize the attributes of the 32 unique scenes.
Therefore thirty-two unique scenes exist, half of which are in a rural setting and the other half in an urban setting. Of the sixteen unique urban and rural scenes, five contain a hopper train consist in the intersection, another five contain a flat train consist in the intersection, another five contain a truck in the intersection, and a final scene exists with no object at all in the intersection.

**Dependent Variable**
After each scene was displayed, a question appeared on the display asking the subject, “What did you see?” The subject could then select train, truck, or nothing using the computer’s mouse. When the subject was finished with the first question, a second question then appeared asking “How confident are you of your previous decision?” The subject could then select anywhere from 0% to 100% confidence in 25% increments. This is the rating question that allows an ROC curve to be developed from a single subject. After the subject answered the second question, she selected a button to continue. A warning queue (a circle at the center of the screen) then appeared for one second followed by the next scene. This cycle of showing a scene followed by questions continued throughout the experiment, and the response to the first and second questions was the dependent variable in this experiment (see Table 2).

<table>
<thead>
<tr>
<th>First Question: What did you see?</th>
<th>Second Question: How confident are you of your decision?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train</td>
<td>Truck or Nothing</td>
</tr>
<tr>
<td>s0</td>
<td>n0</td>
</tr>
<tr>
<td>s25</td>
<td>n25</td>
</tr>
<tr>
<td>s50</td>
<td>n50</td>
</tr>
<tr>
<td>s75</td>
<td>n75</td>
</tr>
<tr>
<td>s100</td>
<td>n100</td>
</tr>
</tbody>
</table>

**Table 2. Experiment I: Dependent Variables**

The subject’s responses of either a truck or nothing were combined into the single response variable \( n \) since we are not interested in distinguishing between the two.
**Summary of Experimental Design**

To summarize, the subjects viewed a total 1200 scenes (the independent variable was described in Table 1). After each scene the subject answered two questions resulting in the dependent variable (see Table 2). Because viewing 1200 scenes and questions takes about three hours, the subject was administered the experiment in two separate sessions to reduce fatigue and boredom. Each session lasted an hour and a half in length, and the second session began approximately 24 hours after the first. The entire experiment was also divided into 20 blocks containing 60 scenes each, where a block is defined as a set of 60 scenes which displays $S_{ijk}$ twenty times, $N_i$ twenty times, and $N_{ijk}$ four times each, where $i = r$ or $u$, $j = h$ or $f$, and $k = 1, 2, 3, 4,$ or $5$ (see Table 3).

<table>
<thead>
<tr>
<th>Block</th>
<th>Setting</th>
<th>Signal Scenes</th>
<th>Noise Scenes</th>
<th>Total number of scenes per block</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rural</td>
<td>20 scenes of $S_{a1}$</td>
<td>20 scenes of $N_r$</td>
<td>60</td>
</tr>
<tr>
<td>2</td>
<td>Rural</td>
<td>20 scenes of $S_{a2}$</td>
<td>4 scenes of $N_1$</td>
<td>60</td>
</tr>
<tr>
<td>3</td>
<td>Rural</td>
<td>20 scenes of $S_{a3}$</td>
<td>4 scenes of $N_2$</td>
<td>60</td>
</tr>
<tr>
<td>4</td>
<td>Rural</td>
<td>20 scenes of $S_{a4}$</td>
<td>4 scenes of $N_3$</td>
<td>60</td>
</tr>
<tr>
<td>5</td>
<td>Rural</td>
<td>20 scenes of $S_{a5}$</td>
<td>4 scenes of $N_4$</td>
<td>60</td>
</tr>
<tr>
<td>6</td>
<td>Rural</td>
<td>20 scenes of $S_{a6}$</td>
<td>4 scenes of $N_5$</td>
<td>60</td>
</tr>
<tr>
<td>7</td>
<td>Rural</td>
<td>20 scenes of $S_{a7}$</td>
<td>4 scenes of $N_6$</td>
<td>60</td>
</tr>
<tr>
<td>8</td>
<td>Rural</td>
<td>20 scenes of $S_{a8}$</td>
<td>4 scenes of $N_7$</td>
<td>60</td>
</tr>
<tr>
<td>9</td>
<td>Rural</td>
<td>20 scenes of $S_{a9}$</td>
<td>4 scenes of $N_8$</td>
<td>60</td>
</tr>
<tr>
<td>10</td>
<td>Rural</td>
<td>20 scenes of $S_{a10}$</td>
<td>4 scenes of $N_9$</td>
<td>60</td>
</tr>
<tr>
<td>11</td>
<td>Rural</td>
<td>20 scenes of $S_{a11}$</td>
<td>4 scenes of $N_{10}$</td>
<td>60</td>
</tr>
<tr>
<td>12</td>
<td>Rural</td>
<td>20 scenes of $S_{a12}$</td>
<td>4 scenes of $N_{11}$</td>
<td>60</td>
</tr>
<tr>
<td>13</td>
<td>Rural</td>
<td>20 scenes of $S_{a13}$</td>
<td>4 scenes of $N_{12}$</td>
<td>60</td>
</tr>
<tr>
<td>14</td>
<td>Rural</td>
<td>20 scenes of $S_{a14}$</td>
<td>4 scenes of $N_{13}$</td>
<td>60</td>
</tr>
<tr>
<td>15</td>
<td>Rural</td>
<td>20 scenes of $S_{a15}$</td>
<td>4 scenes of $N_{14}$</td>
<td>60</td>
</tr>
<tr>
<td>16</td>
<td>Rural</td>
<td>20 scenes of $S_{a16}$</td>
<td>4 scenes of $N_{15}$</td>
<td>60</td>
</tr>
<tr>
<td>17</td>
<td>Rural</td>
<td>20 scenes of $S_{a17}$</td>
<td>4 scenes of $N_{16}$</td>
<td>60</td>
</tr>
<tr>
<td>18</td>
<td>Rural</td>
<td>20 scenes of $S_{a18}$</td>
<td>4 scenes of $N_{17}$</td>
<td>60</td>
</tr>
<tr>
<td>19</td>
<td>Rural</td>
<td>20 scenes of $S_{a19}$</td>
<td>4 scenes of $N_{18}$</td>
<td>60</td>
</tr>
<tr>
<td>20</td>
<td>Rural</td>
<td>20 scenes of $S_{a20}$</td>
<td>4 scenes of $N_{19}$</td>
<td>60</td>
</tr>
<tr>
<td>Total</td>
<td>Rural</td>
<td>400 Signal Scenes per Experiment</td>
<td>800 Noise Scenes per Experiment</td>
<td>1200 Scenes per Experiment</td>
</tr>
</tbody>
</table>

Table 3. Experiment I: Organization

The ordering of the 60 scenes within each block was completely randomized, and the ordering of the blocks during the experiment was randomized as well; however, at no time were scenes
mixed between blocks. The end and beginning of a block was signaled to the subject by a message on the screen followed by a thirty second break.

In terms of SDT, the true state of the world is represented by the independent scene variable in Table 1, whereas the subject’s response is represented by the dependent variable in Table 2. A signal is defined as the independent variable denoted with an S in Table 1, which were scenes with trains in the intersection; and a noise signal is defined as an independent variable denoted with a N in Table 1, which were scenes with trucks or nothing in the intersection. The subject’s response was a signal, $s$, if he or she recognized a train (first column of Table 2) or noise, $n$, if the subject’s response was a truck or nothing (second column of Table 2).

**D. Subjects**

A total of eleven subjects were administered the experiment all of which were licensed drivers with vision better than 20/30. The subject’s ages ranged from 20 to 47 years with a mean of 26 years and standard deviation of 9; 63% of the subjects were female. Eight of the eleven subjects were undergraduates being paid to participate, while the remaining three were volunteers from the Volpe Center.

**E. Instructions and Treatment of Subjects**

The experiment consisted of two sessions not more than 24 hours apart, each lasting approximately one and a half hours. The first session of the experiment began by giving the subject a set of instructions which briefly explained the purpose and format of the experiment (see Appendix A). The vision of the subject was then recorded using a wall mounted vision chart at ambient room light. The subject was then seated in the simulator booth and given a tutorial session that lasted about twenty minutes, which was conducted with no ambient light in order to get the subject’s eyes dark adapted.

In the first part of the tutorial, the subject observed all 32 possible scenes in Table 1 for at least 5 seconds each. The subject was then given the mouse and again shown all 32 possible scenes
followed by the two questions querying the subject’s observation and confidence. When the subject was complete with the practice session, the experiment began. A review of this tutorial was also given at the beginning of the second session, and the second session did not begin until the subject was in the dark for at least fifteen minutes.

During the experiment, the experimenter was in an adjacent room and communicated to the subject via a walkie-talkie. After the subject completed each block of 60 scenes, a message appeared asking the subject to notify the experimenter. The software then gave the subject a mandatory 30 second break and subsequently notified the subject when he or she could resume. It was emphasized to the subject at the beginning of the experiment that it was permitted to take as many breaks as necessary.

2.3 Results

Figure 8 displays the average ROC curves for all eleven subjects; the ROC curves for individual subjects are displayed in Appendix B. The average ROC curves were produced by pooling all of the raw data from the individual subjects and then constructing the ROC curves as if they were one subject.
Figure 8. Average ROC Curves
2.4 Discussion

After viewing these ROC plots, one should notice that both of the urban setting ROC curves are more shifted to the right than the rural setting curves. According to SDT, this would imply that the rural setting had greater detectability. However, by observation of the individual ROC curves located in Appendix B, one can see that subject #10 was responsible for shifting the average ROC curve. Therefore, this shift was simply due to the noise of one subject and does not represent the other ten subjects.

Also note that in all four plots the “no reflectors” curve approximately follows a diagonal from the lower left to the upper right corners of the ROC plot. In signal detection theory, this line is called the line of guessing and indicates that the subject is simply guessing (Green and Swets, 1988). This is expected since the trains with no reflectors are virtually impossible to recognize at night from far distances.

In comparing the four different reflector styles, there is no notable difference in performance. Therefore, one can conclude from this experiment that all four reflector patterns have comparable train recognition qualities.
3. Experiment II

3.1 Objective

The purpose of the second experiment was to determine which of the four reflector patterns described in Figures 1 & 2 gives the best train recognition under a normal driving task on a dark rural road. As the subjects drove a car simulator, they encountered numerous objects (i.e. grade crossings, road intersections with trucks or cars, traffic lights, etc.) and subsequently reported what they recognized. Voice recognition software recorded the subject’s response, and the simulator software determined the distance of recognition.

3.2 Method

A. Apparatus

A driving simulator was created with the purpose of examining the different train reflector patterns. This simulator included a steering wheel, pedal controls, and driving dynamics to simulate a vehicle traveling on a typical two-way American rural road. The road was approximately 40 miles long and contained 22 grade crossings; 20 of which had either flat train cars or hopper train cars passing through. In addition, road markings, speed limit signs, intersections signs, grade crossings signs, and traffic lights were rendered so that this road conformed to U.S. regulations. The posted speed limit was 50mph and most subjects completed the driving task in about an hour. Each grade crossing contained the minimum sign requirements: a railroad warning sign and a crossbuck sign. No other grade crossing features such as flashing lights or reflectorized gates were included; this was done purposefully in order to minimize the conspicuity of the grade crossing. Also, a forest of trees was displayed on both sides of the road at all times so that the subjects were forced to use the scenery directly in front of them to recognize objects.

A sample screen output of the driving simulator is shown in Figure 9. A speedometer was displayed along with the special words the subjects could speak to signal what was seen.
The experimental setup, displayed in Figure 10, was similar to the first experimental setup with addition of a PC to run the voice recognition software, a headset microphone, and a steering wheel/pedal assembly. More information on the software and hardware aspects may be found in Appendix D and E.
B. Experimental Design

In this experiment, subjects drove a car for about one hour along a dark rural road. On this road were many objects including automobiles, lights, signs, trains, and trucks. When the subject recognized any of the aforementioned objects he or she would say the word car, light, sign, train or truck depending on what was recognized. Voice recognition software recorded the subject’s response and compared it with what objects were really in the road (the true state of the world). The distances from these objects when the subject recognized them was also recorded. Although we are not interested in the recognition distances of the objects besides the train, these other objects served as noise and prevented the subject from focusing on just the grade crossings. The subjects viewed 22 grade crossings (20 of which had trains passing through), 40 cars, 20 trucks, and 4 traffic lights in the experiment.
Independent Variable
The independent variable was the type of train reflector pattern on the train and is represented by the variables $H_i$ (hopper train) and $F_i$ (flat train) where $i = 1$ to 5 in Table 4. Each of these variables was displayed to the subject twice; therefore, each subject viewed a total of 20 trains during the experiment. These variables are similar to the scenery variables in Table 1 except that in this experiment only a rural condition existed.

<table>
<thead>
<tr>
<th>Train Type</th>
<th>Reflector Pattern</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hopper Train</td>
<td>Flat Train</td>
</tr>
<tr>
<td>$H_1$</td>
<td>$F_1$</td>
</tr>
<tr>
<td>$H_2$</td>
<td>$F_2$</td>
</tr>
<tr>
<td>$H_3$</td>
<td>$F_3$</td>
</tr>
<tr>
<td>$H_4$</td>
<td>$F_4$</td>
</tr>
<tr>
<td>$H_5$</td>
<td>$F_5$</td>
</tr>
</tbody>
</table>

Table 4. Experiment II: Independent Variables

Dependent Variable
The dependent variable was the distance at which point the subject recognized the train and was measured using voice recognition software. Because there was about a 500 millisecond delay for the voice recognition software to process speech, it is important to realize that errors existed and may have been compounded depending on how fast the driver was driving. However, the abundant data points and experimental randomization ensured that this error was minimized.

Errors
If the subject did not recognize the train or incorrectly recognized the train as another object (i.e. truck or car) an error was recorded. Any recognition distance associated with an error was removed from the data.

C. Subjects
A total of twenty-two subjects, none of which participated in the first experiment, were administered the driving experiment all of which were licensed drivers with vision better than 20/40. The subject’s ages ranged from 18 to 60 years with a mean of 37 years and standard
deviation of 13; 65% of the subjects were male. Seventeen of the twenty-two subjects were volunteers from the Volpe Center and the remaining were undergraduates being paid to participate.

D. Instructions and Treatment of Subjects

The entire experiment for each subject lasted less than two hours. The subject first read instructions (see Appendix C), and then the voice recognition software was trained to his or her voice. This involved the subject saying the words car, light, sign, train, and truck for about two minutes into the headset microphone. The subject then practiced driving for about 7 minutes without the microphone. A practice course allowed the subject to drive in an oval, get accustomed to the controls, and see the various objects present in the experiment. When the subject was comfortable with the driving task, he or she was given the microphone and subsequently practiced driving while speaking the various words. The subject continued this practice until he or she was ready to begin the experiment. During the experiment, the experimenter was in an adjacent room and could communicate with the subject using a walkie-talkie. In addition, the subjects were never told the purpose of the experiment.

3.3 Results

A. Objective Results

Figure 11 plots recognition distance versus reflector pattern; these patterns are described in Table 4. The recognition distances were computed by averaging the data from all twenty-two subjects in the second experiment. The standard deviation is represented by the error bars in Figure 11. An analysis of variance indicated that the means are significantly different (p < .001) with respect to train type (H_i, F_i) and reflector pattern. A Newman-Keuls post-hoc comparison test with respect to the hopper reflector patterns (H_i, i = 1:5) yielded a significant difference (p < .05) between the groups {H_1, H_2, H_4}, {H_3, H_4}, and {H_5} where the group {H_1, H_2, H_4} had the highest recognition distances. However, there was no significant difference within these groups (p > .05). For the flat car, a significant difference (p < .05) existed between the groups {F_2, F_4},
\{F_1, F_3\}, and \{F_5\} where the first group had the highest recognition distances, but there was no significant difference within these groups (p > .05).

![Recognition Distance vs. Reflector Pattern](image1)

**Figure 11. Recognition Distance versus Reflector Type**

**B. Subjective Results**

Figure 12 plots the subjective results of a questionnaire given to the subjects after the driving experiment. In this debriefing, the subjects were asked to rate the reflector patterns according to how well they contributed to train recognition.

![Subjective Recognition Rating vs. Reflector Pattern](image2)

**Figure 12. Subjective Rating versus Reflector Type**
The errors from the driving experiment are compiled in Table 5. The top row of the table indicates what object (truck, car, or nothing) the subject perceived the train as, and the first column indicates the type of train reflector pattern displayed when the error occurred.

<table>
<thead>
<tr>
<th>Hopper Car Errors</th>
<th>Nothing</th>
<th>Car</th>
<th>Truck</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Massed Outline, $H_1$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Vertical Bars, $H_2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Variable Vertical Bars, $H_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Horizontal Bars, $H_4$</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Unreflectorized, $H_5$</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>4</td>
<td>0</td>
<td>5</td>
<td>9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Flat Car Errors</th>
<th>Nothing</th>
<th>Car</th>
<th>Truck</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Massed Outline, $F_1$</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Vertical Bars, $F_2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Variable Vertical Bars, $F_3$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Horizontal Bars, $F_4$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Unreflectorized, $F_5$</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>8</td>
<td>2</td>
<td>4</td>
<td>14</td>
</tr>
</tbody>
</table>

Table 5. Experiment II: Subject Errors
3.4 Discussion

Figure 11 along with the statistical analyses clearly indicates that reflectorized trains were recognized at farther distances than unreflectorized trains, for both the hopper and flat train cars. Furthermore, the Newman-Keuls comparison test shows that the reflector groups \{H_1, H_2, H_4\} and \{F_2, F_4\} performed the best with respect to the hopper and flat train cars. Common to both of these groups is the vertical bar \((i = 2)\) and horizontal bar reflector patterns \((i = 4)\), which both have reflectors distributed along the base of the train car as opposed to the massed outline and variable vertical bars patterns which are lumped towards the ends (see Figures 1 and 2).

However, according to Table 5, three errors occurred with the horizontal bar pattern but none occurred with the vertical bar pattern. Moreover, these errors all involved subjects incorrectly identifying the train as a truck. Since the study by Ford et al. (1996) recommended a standardized train reflector pattern which minimizes confusion with other objects (i.e. truck), these results suggest that the horizontal pattern should not be used.

Interestingly, in the debriefing questionnaire subjects responded favorably with the massed outline reflector pattern on the hopper car but unfavorably with the same pattern on the flat car. This agrees with the experimental results in which a comparison test placed the massed outline pattern in the group with the highest hopper recognition distance and lowest flat train recognition distance. Therefore, the massed outline reflector pattern only performed well with train cars having a substantial height dimension available for marking (i.e. hopper car).

4. Conclusion

The first experiment based on SDT yielded inconclusive results; all four reflector patterns performed equivalently. However, the second experiment based on a realistic nighttime driving simulator concluded that the vertical bar reflector pattern yielded better train recognition performance than the horizontal, massed outline, or variable vertical patterns. Moreover, the second experiment indicates that a distributed pattern has better train recognition qualities than a
lumped pattern on the flat train car. This does not hold true for the hopper car which introduces another dimension (height) available for mounting reflectors.

5. Recommendations

Although a distributed vertical reflector pattern yielded the best performance in this experiment, a larger number of data points will improve accuracy. Therefore, a more comprehensive driving simulator experiment with more subjects should be explored. Furthermore, this study did not take into account any of the mounting issues associated with the different reflector patterns on different possible freight train cars (i.e. flat, hopper, tank, etc.). This will be necessary in order to accurately understand the feasibility and cost of the various reflector patterns on different train cars.
6. References


U.S. Code of Federal Regulations, 49 Part 571.108, Section 5.7.1.4.2


Appendix A. Experiment I Instructions

BACKGROUND

Due to the high number of vehicle accidents at intersections between roadways and railroad tracks (called highway-railroad grade crossings) in the United States, the Department of Transportation is actively examining methods to reduce the number of collisions between trains and motor vehicles. Some of these accidents occur at night and in rural areas where there are no lights and no warning devices that indicate that a train is approaching or already in the grade crossing. Rail cars in the grade crossing may be difficult to see due to a combination of factors that include: a low ambient illumination, a surface that reflects little light back to the motorist. Rail cars are frequently painted in dark colors and accumulate significant levels of dirt. The poor contrast of the rail car compared to the surrounding background makes the motorist’s task of detecting rail cars in the grade crossing a difficult one.

One proposed solution is to add reflective material to rail cars. Reflective materials can reflect back light from the motor vehicle’s headlights and give the driver a much better opportunity to see the rail cars in the grade crossing. The purpose of this experiment is to evaluate the extent to which reflective materials can aid the motorist in identifying the train.

THE EXPERIMENT

You will sit in a simulated motor vehicle that will remain stationary for the duration of the experiment. Located on the screen in front of you will be an intersection. You will see this intersection for a very brief period of time. Your task is to determine whether a vehicle passed through the intersection and what that vehicle is (i.e., a train or a truck).

After each intersection scene is displayed, another display will be appear asking you to indicate which of the following items you saw:

Train  Truck  Nothing

Using the mouse, select the text representing the item you saw. Then you may advance to the next display by selecting the button on the bottom of the screen. The next display asks you to
rate how confident you are of your previous answer. The scale is from 0% to 100%, where 100% means total confidence and 0% means you have no confidence in your previous answer:

0%  25%  50%  75%  100%

Using the mouse, select the percentage that best describes your confidence. When you are ready to proceed to the next intersection scene, click the button on the bottom of the screen.

Scenes are grouped into sessions that will last about ten minutes followed by a short rest period. There are a total of twenty sessions, and you will see a different combination of objects in each session.

If you are uncomfortable or need more time for any reason, let the experimenter know.
Appendix B. Experiment I Individual ROC Curves
Rural - Hopper
Subject 4

Urban - Hopper
Subject 4

Rural - Flat Car
Subject 4

Urban - Flat Car
Subject 4
Appendix C. Experiment II Instructions

Driving Experiment

The purpose of this experiment is to study the ability of automobile drivers to recognize objects on a rural road at night.

This experiment will last about an hour and involves driving a simulated motor vehicle using a steering wheel and gas/brake pedals (no clutch, simulating an automatic transmission). Your task is to identify the objects you see in the roadway as you drive the simulated motor vehicle. As you drive, follow the posted speed limit and slow down when appropriate (at intersections, etc.). Where it is necessary to turn left or right at an intersection, an arrow will indicate which direction.

During the experiment, you will see a series of objects like those you would encounter on roadways in the real world. As soon as you can identify the object, say what the object is by saying the name of the object aloud. To simplify the identification task, use one of the five words below:

Car  Sign  Light  Train  Truck

For example, if you are driving along and you see another car, saying the word car, if you see any sort of sign on the roadway you will say sign, if you see a traffic light you will say light, if you see a train (at a grade crossing) you will say train, and if you see a truck you will say truck.

If for some reason you need to temporarily stop the experiment, saying the word pause will pause the experiment and allow you to stretch and/or contact the experimenter. You do not need to memorize these words as they will displayed below the out-the-window view during the entire experiment. When you are ready to continue say the word resume.
Before the experiment begins, you will train the voice recognition software to recognize your voice. It is important that you speak normally, as if you were talking to someone only a few feet away in a quiet room. Not only is this the optimal loudness for the software, but it will also keep you from getting tired since the experiment lasts an hour. As you might expect, slurring your speech will impede voice recognition and, therefore, it is important that you speak in a consistent and clear manner.

You will be in voice contact with the Experimenter at all times in the adjacent room via a walkie-talkie. If at anytime you would like to speak to the administrator then simply say pause and the driving simulator will stop. When you are ready to resume, simply say resume and the simulator will resume. These words will be displayed at all times.
Appendix D. Hardware

Driving Interface
The driving interface used in Experiment II was the Thomas Super Wheel manufactured by Thomas Enterprises (see Figure 13). At a cost of approximately $350, this metal framed steering wheel and pedal assembly was more robust than cheaper off-the-shelf plastic models found at most computer stores. Although this steering wheel was originally designed to be used with a PC’s joystick port, simple rewiring made it adaptable to the CerealBox analog/digital converter.

Figure 13. Picture of the Steering Wheel/Pedal Assembly

The steering wheel contained one potentiometer and the pedal assembly contained two potentiometers which were connected to the CerealBox analog/digital converter as shown in Figure 14.
CerealBox: An Analog/Digital Converter
Manufactured by BG Systems, the CerealBox is an analog/digital converter specifically designed for using the serial port of UNIX workstations. The model used in this experiment was the LV824-G which offers up to 8 analog inputs, 24 digital inputs/outputs, 3 analog outputs, and a +5Vdc output. For this experiment, only 2 of the 8 analog inputs were required along with the +5Vdc output. Because the normal driving task does not require pressing the brake and accelerator pedal at the same time, these two pedals shared a single analog channel which the simulator software subsequently separated.

Computers: SGI Indigo2, Pentium PC
A Silicon Graphics (SGI) Indigo2 workstation served as the graphics engine for both experiments, where the screen output of the SGI was sent to a Barco 808S projector for viewing.
on a screen. A Pentium PC was equipped with the features necessary to run the voice recognition software including 32MB RAM, an Intel Pentium processor, and a SoundBlaster 16bit sound card. Although sound generation was developed to simulate the vehicle’s engine, it could not be used because the voice recognition software required full use of the sound card. Installing a more expensive duplexing sound card could allow both sound generation and voice recognition at the same time.
Appendix E. Software

Software Overview
Both experiments utilized software on a Silicon Graphics (SGI) workstation to render the various scenery and simulate driving dynamics. In addition, the second experiment required use of PC based voice recognition (VR) software to input the response of subjects. Because the VR software was a commercial package intended for dictating speech, its output was keystrokes spelling the various words being spoken. To get the subject’s response recorded on the SGI, an interface application had to be developed to receive the output of the VR and send it over a network to the SGI (see Figure 15).

![Diagram of software overview](Image)

Figure 15. Overview of Software

Voice Recognition Software
The voice recognition software used in this experiment was DragonDictate developed by Dragon Systems of Newton, MA. Intended as a large vocabulary system for text dictation, this commercial software package was modified to behave as small vocabulary system through hacking on a few of the settings. During normal use, the software has many modes such as Dictate, Mouse Movement, and Sleep. In sleep mode, the software ignores all speech through the microphone except for the one phrase “Wake Up”. When Dragon Dictate recognizes this phrase it will enter another mode. Using the Vocabulary Manager of DragonDictate, the phrase “Wake Up” was removed and replaced with the seven words used in this experiment: car, sign, light, train, truck, pause, and resume so that the software would only search for these words and nothing else. If the subjects did not enunciate one of the seven words correctly, the software would output nothing. Thus, it behaved as a small vocabulary system.
**Simulator Software**

The simulator software was written in C++ using the OpenGL application programming interface (API) and debugged using the DBX debugger. C++ was the programming language of choice because it was the most prominent object-oriented language at the time of this research. A procedural language such as C, albeit faster, was rejected due to the overwhelming benefits of Object Oriented Programming that C++ offers. On the other hand, a more high level language such as Java was rejected due its large overhead.

The OpenGL API was selected because it was the most prominent graphics libraries at the time of this research. Other graphics libraries such as IrisGL, Direct3D, or Farenheit were rejected because they were either no longer supported, not SGI-UNIX compatible, or still under development, respectively. The biggest benefit OpenGL offered was platform independence, which means simulator software developed on an SGI can be ported to a PC if desired. In addition, all windows event handling was done using the Graphics Utility Toolkit (GLUT) developed by SGI. This toolkit maintains pure platform portability for all windowing events.

The directory structure of the simulator is depicted in Figure 16. The directory *Carsim* contains all the C++ source code pertinent to the simulator. The directory *CerealBox* contains the driver software developed by BG Systems for using the CerealBox analog/digital converter. The directory *RGBAFiles* contains the picture files used in the simulator for objects such as the street signs. The directory *Network* contains the ethernet software necessary for networking with the PC. This software was developed by former Human-Machine Systems Laboratory members and is described in depth in “Sound Enhancements for Graphical Simulations” by Steven Villareal (1997). The *GLUT* directory contains the source code for the Graphics Library Toolkit developed by Mark Kilgard formerly of Silicon Graphics. GLUT was used as the windows and event management.
Interface Software

The interface software was a simple C++ program that inputted keystrokes from the VR software and then sent the data to the SGI through a network. The PC network software was developed by Steven Villareal which is discussed in his thesis (1997).