

# An Application of Signal Detection Theory for Understanding Driver Behavior at Highway-Rail Grade Crossings

Michelle Yeh and Jordan Multer  
United States Department of Transportation  
Volpe National Transportation Systems Center  
Cambridge, MA

Thomas Raslear  
Federal Railroad Administration  
Washington, DC

We used signal detection theory to examine if grade crossing warning devices were effective because they increased drivers' sensitivity to a train's approach or because they encouraged drivers to stop. We estimated  $d'$  and  $\beta$  for eight warning devices using 2006 data from the Federal Railroad Administration's Highway-Rail Grade Crossing Accident/Incident database and Highway-Rail Crossing Inventory. We also calculated a measure of warning device effectiveness by comparing the maximum likelihood of an accident at a grade crossing with its observed probability. The 2006 results were compared to an earlier analysis of 1986 data. The collective findings indicate that grade crossing warning devices are effective because they encourage drivers to stop. Warning device effectiveness improved over the years, as drivers behaved more conservatively. Sensitivity also increased. The current model is descriptive, but it provides a framework for understanding driver decision-making at grade crossings and for examining the impact of proposed countermeasures.

## INTRODUCTION

The Federal Railroad Administration (FRA) needs a better understanding of driver decision-making at highway-rail grade crossings. Grade crossing safety has improved; from 1994 through 2003, the number of grade crossing accidents decreased by 41 percent and the number of fatalities fell by 48 percent. However, many of the accidents that occurred were the result of driver error or poor judgment (Office of the Inspector General, 2004). Thus, we were interested in modeling drivers' decision-making strategies to understand the factors that contribute to decisions at grade crossings and to establish a framework for research to evaluate the impact of proposed countermeasures.

We applied signal detection theory (SDT) to describe drivers' decisions whether to stop or proceed. In this application of SDT, the train is the *signal*, and it provides visual and auditory cues as to its approach (e.g., alerting lights and the sound of the horn). Other information at the grade crossing create *noise* that may compete with the signal, such as the flashing lights, gates, and bells at some crossings or sounds from inside the vehicle.

There are two possible states of the world when a driver is approaching a grade crossing: that a train is approaching or that it is not. The driver must then make one of two choices: stop or proceed. A driver's action at a grade crossing can thus be described by the 2 x 2 signal-response matrix in Table 1. Compliant behavior at grade crossings is indicated by the highlighted cells. A **valid stop** is the decision to stop when a train is approaching (a *hit* in SDT), and a **correct crossing** describes the decision to proceed when a train is not approaching (a *correct rejection*). Of concern is a *miss* at a grade crossing—the decision to proceed when a train is close, which can result in an **accident**. Finally, a **false stop**, or *false alarm*, in the grade crossing situation results when a driver stops when no train is approaching.

		State of the World	
		Train is close	Train is not close
Driver Response	Yes (Stop)	<b>Valid Stop</b> (driver stops at crossing)	<b>False Stop</b> (driver stops unnecessarily)
	No (Proceed)	<b>Accident</b> (driver does not stop)	<b>Correct Crossing</b> (driver safely crosses tracks)

Table 1. Signal-Response Matrix for a Driver at a Grade Crossing.

Raslear (1996) applied this framework of driver decision-making to examine the effectiveness of eight grade crossing warning devices and to determine if the source of their effectiveness was because the warning devices increased the signal-to-noise ratio at the grade crossing (*sensitivity*) or because they encouraged drivers to stop (*bias*). The grade crossing systems consisted of four passive warning devices (no protection, the crossbuck, stop sign, or other signs or signals), classified as such because the information they provide does not change regardless of the presence or approach of a train, and four active warning devices (gates, flashing lights, highway traffic signals, and special warning devices). Device effectiveness was calculated for each warning device as a proportion of the accident risk at a grade crossing relative to the observed probability of an accident. The data used for the analysis were collected in 1986.

The results indicated that active grade crossing warning devices were more effective than passive warning devices and that the source of this effectiveness was primarily the result of how they affected drivers' bias. Active warning devices encouraged drivers to stop with gates having the highest bias.

Grade crossings with no protection were the least effective, and drivers were more likely to proceed than to stop when encountering one of these crossings. The warning devices showed little difference in sensitivity, which was relatively high overall, suggesting that the train generally presents a tremendous signal relative to the background noise.

More than 20 years later, safety at grade crossings has improved significantly, and it was of interest to examine if these improvements could be captured using the SDT framework. Thus, the purpose of this research was to update the analysis conducted by Raslear and to compare our findings to his to understand how driver decision-making at grade crossings had changed.

## METHOD

We estimated sensitivity, bias, and device effectiveness for eight grade crossing warning devices:

- No protection,
- Other signs or signals,
- Crossbuck,
- Stop sign,
- Special warning devices,
- Other activated warning devices (e.g., highway traffic signals, wigwags),
- Flashing lights, and
- Gates.

The first four warning devices listed are considered passive warning devices, and the last four are active warning devices.

Sensitivity was estimated using  $d'$  and calculated as:

$$d' = z[P(VS)] - z[P(FS)].$$

In the formula,  $P(VS)$  is the probability of a valid stop (a hit), and  $P(FS)$  is the probability of a false stop (a false alarm). Bias was estimated using  $\beta$  and calculated as:

$$\beta = \frac{y(VS)}{y(FS)}, \text{ where}$$

$$y(VS) = 0.3989e^{-z[P(VS)]^2/2}, \text{ and}$$

$$y(FS) = 0.3989e^{-z[P(FS)]^2/2}$$

The values for  $P(VS)$  and  $P(FS)$  can be estimated from accident data. An overview of the calculations is presented here. Additional details and formulas are provided in Raslear (1996).

By definition,  $P(VS)$  is equal to 1 minus the probability of an accident (i.e.,  $1 - P(AC)$ ).  $P(AC)$  can be calculated from the accident rate for each warning device, equalized for exposure. Each warning device is used at a different number of crossings, and each of these crossings is frequented by a different number of trains per day and a different number of cars per day. Grade crossing warning devices are selected as a function of the number of trains and cars per day, so active warning devices will have higher exposure than passive warning devices. To take all these factors into consideration when determining the probability of an accident,  $P(AC)$  was estimated for each grade crossing warning device as the accident rate per crossing per train per highway vehicle per minute.

$P(FS)$  was not as easy to define as  $P(VS)$ . Rather, it was more straightforward to estimate the probability of a correct crossing,  $P(CC)$ , and to calculate  $P(FS)$  as 1 minus  $P(CC)$ .  $P(CC)$  reflects the probability that a car and a train will not have an accident if a car does not stop.  $P(CC)$  can be estimated as 1 minus the maximum accident risk at a grade crossing,  $P(AC)_{\max}$ , where  $P(AC)_{\max}$  is the probability that a train and a car are at a crossing simultaneously and that neither can stop. Therefore:

$P(FS) = 1 - P(CC) = 1 - [1 - P(AC)_{\max}] = P(AC)_{\max}$   
 $P(FS)$  is equal to the accident risk associated with each warning device.

$P(AC)_{\max}$  was calculated as the product of the probability that one or more trains would be observed at a grade crossing in a 1-minute period ( $p_T$ ) and the probability that one or more highway vehicles would be observed at a grade crossing in a 1-minute period ( $p_H$ ). The values for  $p_T$  and  $p_H$  can be estimated from the train rate per day at a crossing and the average annual daily traffic (AADT) at a crossing, respectively. Note that  $p_T$  and  $p_H$  are also used to estimate exposure for  $P(AC)$ .

Finally, we also wanted to obtain a measure of device effectiveness. Simply examining the accident rate can not provide such a measure because the accident rate speaks only to the observed frequency of accidents and not to the accident risk, which is the number of accidents that would have occurred if no warning device was present. Thus, to estimate device effectiveness, Raslear (1996) computed the ratio between the maximum probability of an accident (the accident risk) and the observed probability of an accident (the accident rate). That is:

$$\text{Device Effectiveness} = \frac{P(FS)}{P(AC)}$$

The information needed for this analysis was contained in two separate FRA databases: the Highway-Rail Grade Crossing Accident/Incident database and the Highway-Rail Crossing Inventory. The Highway-Rail Grade Crossing Accident/Incident database records details on each accident occurring at a highway-rail grade crossing per calendar year, and it was used to calculate the total number of accidents for each of the eight grade crossing warning devices. The Highway-Rail Crossing Inventory contains a list of all grade crossings, and it was used to determine the number of public grade crossings protected by each of the eight warning devices, the median number of trains per day for the crossings, and the median AADT at the crossings.

Data from the Highway-Rail Grade Crossing Accident/Incident database can be downloaded from the FRA Office of Safety web site for a specific year, but the Highway-Rail Crossing Inventory does not provide a similar yearly snapshot. Information from the databases is summarized in FRA's Annual Reports (e.g., see FRA (2008)), so data for the most recent year available—2006—were used for this analysis. Data from 1986, the year for which Raslear (1996) conducted his analysis, were also examined as part of this effort for comparison.

One problem with the data in the Highway-Rail Crossing Inventory is that the fields for trains per day and AADT are

not updated each year. In fact, a comparison of the median AADT values recorded in 1986 to that recorded in 2006 showed that AADT *decreased* for some warning devices. This finding was contrary to other data that showed an overall increase in the number of highway vehicles and national vehicle miles travelled (VMT), which would suggest a corresponding increase in AADT at grade crossings (see the Federal Highway Administration's (FHWA) *Highway Statistics* series). Consequently, we were concerned with the reliability of the measures we used to equalize for exposure, and we sought ways to adjust our estimates for the train rate and highway vehicle rate per day at grade crossings.

We used information regarding the increase in the overall number of trains from 1986 to 2006 to capture changes in the frequency of trains at a grade crossing. The number of trains was calculated by dividing the number of train miles traveled with the number of track miles for each year from 1986 through 2006, using data from the FRA and the American Association of Railroads. The 1986 rate for trains per day from the Highway-Rail Crossing Inventory was not changed. Instead, the train rate per day for each warning device in 2006 was estimated from the 1986 data and increased proportionally to reflect the change in the number of trains since 1986. We expected that this calculation would reflect the changes in exposure while minimizing the effect of any error in the train rate per day reported in the Highway-Rail Grade Crossing Inventory.

A similar process was conducted to estimate in the highway vehicle rate. We collected information on national VMT from 1986 through 2006, as reported by FHWA. The AADT for 1986 was unchanged from that reported in the Highway-Rail Crossing Inventory, but the AADT values used for 2006 were calculated using the 1986 data as a baseline and increased proportionally to reflect the changes in VMT since 1986.

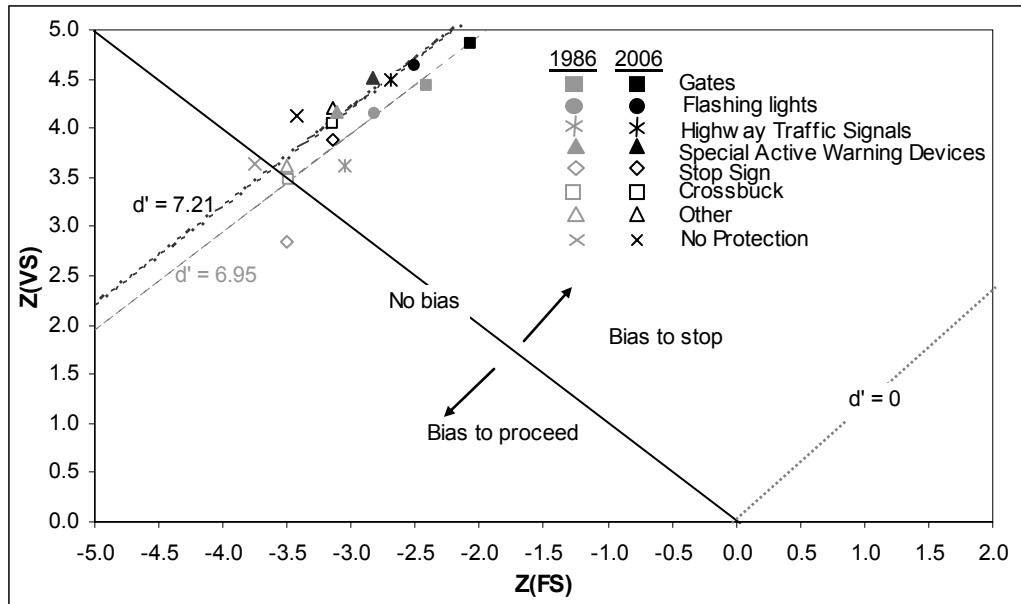
## RESULTS

In 1986, 5,710 accidents occurred at 192,454 grade crossings (approximately 0.030 accidents/crossing). The number of accidents dropped by over 50 percent by 2006, with 2,387 accidents at 139,886 grade crossings (approximately 0.017 accidents per crossing). During the same time period, exposure increased; the total trains per day at a grade crossing increased by 125 percent, and the AADT by 64 percent. This data was used to develop estimations for  $d'$ ,  $\beta$ , and device effectiveness for 1986 and 2006, as shown in Table 2. (Note that a direct comparison of these estimations to the overall change in the accident rate in the same time period is difficult because the estimations reflect probabilities.) Figure 1 shows a transformed Receiver-Operator Characteristic (ROC) curve for the data sets using a plot of  $z(VS)$  versus  $z(FS)$ . Data for 1986 are drawn in gray in the figure and data for 2006 in black.

In first examining sensitivity, estimations of  $d'$  were fairly high for both years (6.95 in 1986 and 7.21 in 2006), suggesting that the train presents a fairly salient signal at a grade crossing. Proportionally, the change in sensitivity throughout the 20-year period was approximately 4 percent for all the warning devices combined with the greatest individual increase observed for the stop sign. This change, although small, was significant,  $t(7) = -3.16$ ,  $p < 0.05$ . The mean  $d'$  for each year are represented by the dotted lines in Figure 1; note that because the figure shows  $z(VS)$  versus  $z(FS)$  rather than  $P(VS)$  versus  $P(FS)$ , the sensitivity contours have a slope of 1. As the figure shows, the data points for each warning device for a given year fall fairly close to the mean for that year, thus indicating that little difference was in the signal-to-noise ratio across the warning devices. In 2006, sensitivity ranged from a low of 6.92 (gates) to a high of 7.54 (no signs or signals). The pattern of results makes sense intuitively; grade crossings with no signs or warning signals generate less perceptual "noise" that compete for the driver's attention, so the signal-to-noise ratio may be greater and the train more salient.

Warning device	1986			2006		
	$d'$	$\beta$	Device Effectiveness	$d'$	$\beta$	Device Effectiveness
No Signs or Signals	7.40	1.49	0.66	7.54	0.07	16.41
Other Signs or Signals	6.98	1.03	0.97	7.35	0.02	61.09
Crossbuck	7.12	0.67	1.53	7.20	0.04	32.99
Stop Signs	6.34	8.24	0.10	7.03	0.07	16.26
Special Active Warning Devices	7.28	0.02	58.10	7.34	0.002	678.34
Highway Traffic Signals, Wigwags, Bells, or Other Activated Warning Device	6.68	0.15	7.75	7.19	0.002	1065.13
Flashing Lights	6.96	0.01	141.18	7.14	0.001	3229.16
Gates	6.84	0.001	1667.92	6.92	6.49E-05	31950.29
<b>AVERAGE</b>	<b>6.95</b>	<b>1.45</b>	<b>234.77</b>	<b>7.21</b>	<b>0.03</b>	<b>4629.11</b>

Table 2. Revised  $d'$ , beta, and device effectiveness values for 1986 and 2006.



**Figure 1. ROC plot: Adjusted values for 1986 and 2007.**

Although grade crossings with no warning devices have a high signal-to-noise ratio, the lack of information about the presence of a crossing and whether or not a train is approaching tends to encourage drivers to proceed through the crossing, as shown by the estimations for  $\beta$ . A value of  $\beta$  equal to 1 represents no bias (in Figure 1, this is depicted as the solid diagonal extending from the top left to point (0, 0)), a value of  $\beta$  less than 1 indicates a willingness to stop (the area above the diagonal), and a value of  $\beta$  greater than 1 indicates an inclination to proceed (the area below the diagonal). The average  $\beta$  in 1986 was 1.45, indicating that drivers were more likely to proceed than to stop, but much of this risky behavior reflected decisions at passive grade crossings. By 2006, drivers were more likely to stop than proceed at a grade crossing; the estimations of  $\beta$  had improved from 1986 values by almost 50 times to an average of 0.03. A comparison of  $\beta$  across the two years, using the natural logarithm, indicated that this shift was significant,  $t(7) = 10.62$ ,  $p < 0.05$ . This conservative change in driver decision-making was observed for all the warning devices (in Figure 1, the greater willingness to stop can be seen in a shift up and to the right for each of the data points from 1986, and all the data points fall above the solid diagonal where  $\beta = 1$ ). In both years examined in this analysis, the most conservative decisions were made at grade crossings protected with gates. This finding is not surprising, because gates are the most salient active protection system. In 2006, the riskiest behavior occurred at grade crossings with no signs or signals and those that were protected with stop signs, but even at these grade crossings, drivers exhibited a willingness to stop ( $\beta = 0.07$ ).

The changes in  $d'$  and  $\beta$  can be considered with respect to the device effectiveness ratios shown in Table 2. The higher the ratio is above 1, the greater the effectiveness of the warning device. A ratio less than or equal to 1 indicates that a warning device is not effective because the accident rate is higher than the accident risk. As the data in Table 2 indicate, device effectiveness improved considerably from 1986 to

2006. As the accident risk increased, there was no corresponding increase in the accident rate. Active warning devices (the last four warning devices listed in the table) were in general more effective than passive warning devices (the first four warning devices in the table). Gates were the most effective warning device, with a device effectiveness ratio 10 times greater than the second most effective warning device, flashing lights. For both years of the analysis, the stop sign was the least effective warning device, although its effectiveness had greatly improved in 2006 (16.26) relative to 1986 (0.10), when there appeared to be a greater likelihood of accidents than would be expected given the train and highway vehicle rates.

We were interested in determining the source of the effectiveness of the warning devices to determine if changes in the signal-to-noise ratio or changes in driver's bias accounted for the increased device effectiveness. Correlations were performed between  $d'$  and device effectiveness and between  $\beta$  and device effectiveness using the 2006 data. For the purposes of this analysis, device effectiveness was transformed using the logarithmic scale and  $\beta$  was transformed using the natural logarithm. We found a statistically significant relationship between device effectiveness and  $\beta$  ( $r = -0.99$ ,  $p < 0.05$ ), suggesting that the setting of the response criterion accounted for most of the variance in device effectiveness. Changing the signal-to-noise ratio had little effect on device effectiveness; the correlation between device effectiveness and  $d'$  was not significant ( $r = -0.55$ ,  $p > 0.05$ ). The results are consistent with those reported by Raslear (1996), who concluded that the source of the effectiveness of grade crossing warning devices appears to be in how they influence the setting of bias in the decision-making process.

## DISCUSSION

The results of the SDT analysis provide a glimpse into how different warning devices influence drivers' decisions at

grade crossings. The analysis showed that in the 20-year period examined, drivers became more conservative in their decisions at grade crossings, and this inclination to stop at grade crossings played a large role in improving the effectiveness of all warning devices by an order of magnitude. Sensitivity of the warning devices at grade crossings was high in 1986, and it increased only slightly but significantly through 2006. Thus, this change, although small, also contributed to reducing the accident risk at grade crossings, even though it was not reflected in device effectiveness. This finding makes sense intuitively; the warning devices are located at the grade crossing and are not part of the train, so they can not enhance the signal by themselves. More importantly, the results of the analysis also suggest that the warning devices did not contribute to increasing the noise at grade crossings.

The estimations of  $d'$  and  $\beta$  also provide some insight into driver behavior at active versus passive grade crossings. The values of  $d'$  estimated for 2006 were only slightly higher at passive crossings relative to active grade crossings (7.28 vs. 7.14), suggesting that the signal-to-noise ratio was fairly similar and the approach of a train was equally detectable. However, there was a much greater willingness to stop at active crossings ( $\beta = 0.001$ ) than at passive crossings ( $\beta = 0.05$ ). This behavior contributed to a device effectiveness ratio almost 300 times greater for active warning devices than passive warning devices. Unfortunately, the model can not speak to the drivers' motivations for stopping or proceeding. However, the pattern of results is consistent with the hypothesis that low train frequencies, which are more typical of passive crossings than active crossings, predispose drivers not to expect a train and thus biases them in general to proceed.

Of the eight warning devices, the stop sign had the lowest estimate for  $d'$  and one of the highest for  $\beta$ . In other words, it was the least sensitive, but at these crossings, drivers made the most risky decisions. The use of a stop sign at grade crossings is controversial; the similarities between a driver's decision at a grade crossing and that at a highway intersection initially prompted interest in its use, but the results of several observational studies have reported low rates of compliance (see Yeh and Multer, 2008). In fact, the estimate of  $\beta$  for 1986 was 8.24, reflecting an inclination to proceed at stop sign-protected crossings. Even though compliance improved by 2006, the device effectiveness ratio was still almost equal to that of grade crossings with no warning device at all.

Another concern with the use of the stop sign is that a driver, stopped at the crossing, may not be able to effectively judge the speed of an approaching train, because cues regarding the lateral movement of the train are not as available. The estimates for sensitivity may provide some evidence in this regard; if drivers were less able to detect cues of an approaching train when stopped at the crossing, we would expect to see a lower signal-to-noise ratio at these grade crossings than for those with warning devices that required drivers to slow on the approach (e.g., the crossbuck). Indeed, this is the case; estimates of  $d'$  for the stop sign in the two

years examined were generally lower than the other passive warning devices.

Thus, the current model can be used to provide insight into driver decision-making at grade crossings. Although descriptive in nature, the model can be used in conjunction with field studies or laboratory experiments to contribute to an understanding of driver behavior. Our next step is to apply the model to examine and quantify the impact of different countermeasures that have improved safety at grade crossings. Countermeasures can be classified into two categories: those that increase the detectability of the train (sensitivity), such as the use of alerting lights or reflecting train cars, and those that encourage drivers to stop (bias), such as imposing tangible penalties for grade crossing violations. We are particularly interested in trying to parse out the individual effects of the different countermeasures to examine the corresponding changes in sensitivity and bias. The results will be used to refine the SDT model, so it can be better applied to meet a variety of needs to understand and improve driver behavior.

## ACKNOWLEDGEMENTS

This research was conducted with funding from the FRA's Office of Research and Development. The views expressed herein are those of the authors and do not necessarily reflect the views of the John A. Volpe National Transportation Systems Center, the Research and Innovative Technology Administration, or the United States Department of Transportation.

## REFERENCES

- Federal Highway Administration. (2007). *Highway Statistics 2006*. Washington, D.C.: Federal Highway Administration. Online at: <http://www.fhwa.dot.gov/policy/ohpi/hss/index.cfm>
- Federal Railroad Administration. (2008). *Railroad Safety Statistics: 2006 Annual Report*. Washington, D.C.: Federal Railroad Administration. Online at: <http://safetydata.fra.dot.gov/officeofsafety/>
- Office of the Inspector General (2004). *2004 Audit of the Highway-Rail Grade Crossing Safety Program* (Report Number: MH-2004-065). Washington, D.C.: Federal Railroad Administration, Office of the Inspector General.
- Raslear, T.G. (1996). Driver behavior at rail-highway grade crossings: A signal detection theory analysis. In A. A. Carroll and J. L. Helser (Eds.), *Safety of highway-railroad grade crossings*. Research needs workshop. Volume II – Appendices (Report No. DOT/FRA/ORD-95/14.2, DOT-VNTSC-FRA-95-12.2, pp. F9-F56). Washington, DC: U.S. Department of Transportation.
- Yeh, M. and Multer, J. (2008). *Driver Behavior at Highway-Railroad Grade Crossings: A Literature Review from 1990–2006* (Report Nos. DOT-VNTSC-FRA-08-03 & DOT/FRA/ORD-08/03). Cambridge, MA: U.S. DOT Volpe Center.