

Evaluating the Impact of Grade Crossing Safety Factors through Signal Detection Theory

Michelle Yeh
Federal Aviation Administration
Washington, DC

Thomas Raslear
Federal Railroad Administration
Washington, DC

Jordan Multer
Volpe National Transportation Systems
Center
Cambridge, MA

The purpose of this effort was to apply signal detection theory to descriptively model the impact of five grade crossing safety factors to understand their effect on driver decision making. The safety factors consisted of: improving commercial motor vehicle driver safety through federal regulations, increasing locomotive conspicuity with alerting lights, increasing locomotive conspicuity with reflectors, increasing sight lines, and improving warning device reliability. We estimated d' and β for eight warning devices associated with each safety factor using data from the Federal Railroad Administration's Highway-Rail Grade Crossing Accident/Incident database and Highway-Rail Crossing Inventory. We also calculated ω^2 for each safety factor and device type to examine the reliability of each independent variable on grade crossing safety. The analysis indicated that the first four safety factors listed above were generally equally effective in improving grade crossing safety. Warning device reliability (which pertained to active warning devices only) still contributed to improvements in grade crossing safety, but the effects were more muted. Grade crossing devices (and particularly active warning devices) are an important safety tool because they increase drivers' inclinations to stop, and this bias to stop has a greater impact than improving the driver's ability to detect the train.

INTRODUCTION

Driver error or poor judgment continues to play a significant role in grade crossing accidents (Office of the Inspector General, 2004). Traditionally, improvements in safety at grade crossing focused on the "Three E" approach: education, enforcement, and engineering. The Three E's do not directly address drivers' decisions and motivations that lead to a grade crossing violation, however, although these are important considerations. Therefore, to better understand driver behavior at grade crossing, Raslear (1996) used signal detection theory (SDT) to model driver decision-making strategies in an attempt to understand the motivational factors that contribute to grade crossing violations.

In this application of SDT, the train is the *signal*; it provides visual and auditory cues as to its approach or presence at a grade crossing (e.g., alerting lights, reflectors). Other information at the grade crossing create *noise* that may compete with the signal, such as the flashing lights and gates at some crossings or sounds from inside the vehicle. The signal-response matrix to describe a driver's action at a grade crossing is presented in Table 1. As the table shows, there are two possible states of the world when a driver approaches a grade crossing: a train approaches or it does not. The driver must then make one of two choices: stop or proceed. Compliant behavior at grade crossings is indicated by the highlighted text. 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)	Valid Stop (driver stops at crossing)	False Stop (driver stops unnecessarily)
	No (Proceed)	Accident (driver does not stop)	Correct Crossing (driver safely crosses tracks)

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

In one application of this model, Raslear examined the effectiveness of eight grade crossing warning devices to determine whether the warning devices were effective because they increased the signal-to-noise ratio at the grade crossing (*sensitivity*) or because they encouraged drivers to stop (*bias*). Using data from 1986, Raslear reported that the analysis suggested that grade crossing warning devices are effective because they encourage drivers to stop. Yeh, Multer, and Raslear (2009) updated this analysis and compared the findings to more recent data (2006) with similar results. Additionally, the results indicated that warning device effectiveness improved over the 20-year period examined, as drivers behaved more conservatively.

The results by Raslear (1996) and Yeh, et al. (2009) demonstrate the improvements in safety at grade crossings, but the analysis does not speak to what led to this improvement. The number of grade crossing accidents decreased by 41% between 1994 and 2003 (Office of the Inspector General, 2004), a time period in which a number of safety measures were implemented. The Federal Railroad Administration (FRA) was interested in measuring the impact of the different programs. Thus, the purpose of this research effort was to extend the SDT model to establish a framework for research for evaluating the impact of proposed countermeasures.

In 2008, Horton, et al. quantified the benefits of the different safety programs by measuring the reduction in incidents attributable to each program. Horton, et al. identified factors considered to be a “success” in improving grade crossing safety through a series of literature reviews and stakeholder/industry group discussions. Next, they developed an estimate of “success” by comparing the percentage of incidents from 1994 through 2003 that could be attributed to a factor in each year and measuring the percent reduction in incidents. An incident was classified as being attributable to one success factor, more than one of the success factors, or none of the factors by using information provided in the data fields of the grade crossing accident reports recorded in the FRA Highway-Rail Grade Crossing Accident/Incident database. The impact and percent reduction for each of the success factors was then calculated using a count of those incidents attributable to that one success factor only. Table 2 presents the success factors and the estimated percent reduction of incidents. As the table shows, two factors, “improving CMV driver safety through federal regulations” and “increasing locomotive conspicuity through the use of alerting lights, accounted for approximately 48% of the reduction in accidents. The other three factors each accounted for a 3% to 4% change in the 10-year period examined. The cumulative percent reduction for all five factors totaled 58%.

Safety Factor	Reduction
1. Improving CMV driver safety through federal regulations	34.6%
2. Increasing locomotive conspicuity with alerting lights	13.6%
3. Increasing sight lines	3.6%
4. Improving warning device reliability	3.1%
5. Improving the reliability of motor vehicles	3.1%

Table 2. Percent reduction of incidents for each isolated "success" factor.

We wanted to use the same approach developed by Horton, et al. but focus on the impact of those programs in terms of their effect on driver decision making. We expected that Factors 2 and 3 would exert a primary influence on sensitivity (i.e., the detectability of the train) and Factors 1 and 4 would influence bias. We hypothesized that Factor 5 would not exert an influence on driver decision-making because it

was measured on the basis of whether an accident involved a car that broke down or stalled on the highway tracks, so we excluded it from analysis via the SDT framework. Instead, we included the use of reflectors to increase locomotive conspicuity as another factor, since this was expected to influence sensitivity, even if it had not been included in Horton, et al. Note that although we distinguish the safety factors according to how they were expected to influence sensitivity and bias, we did not expect that the two measures would change independently. Rather, we expected that a change in one could have an observable effect on the other.

METHOD

We estimated d' and β to descriptively measure the effect of five safety factors on driver decision making. We used the definitions developed by Horton, et al. to calculate the number of accidents associated with Safety Factors 1-4 (as identified in Table 2). Our fifth safety factor was improving locomotive conspicuity through the use of reflectors. That safety factor was not included in the analysis by Horton, et al., but had been defined in previous research sponsored by the FRA (see Carroll, et al., 1999). Thus, the number of accidents for each safety factors was calculated as follows:

- *Improving CMV driver safety through federal regulations:* The accidents associated with this safety factor resulted from a motor vehicle that would require a Commercial Drivers' Licenses (CDL).
- *Increasing locomotive conspicuity with alerting lights:* This safety factor was defined by accidents in which rail equipment struck a motor vehicle either at dusk, dawn, or dark, because the use of alerting lights provides more benefit at night than during the day.
- *Increasing sight lines:* Accidents associated with this safety factor noted one of five visual obstructions at the grade crossing: a permanent structure, standing railroad equipment, topography, vegetation, and other.
- *Improving warning device reliability:* This safety factor pertained only to grade crossings with active warning devices. We identified accidents that resulted because the warning signal malfunctioned; the malfunctions consisted of alleged and confirmed warning times that were too long or too short or when no warning was presented.
- *Increasing locomotive conspicuity with reflectors:* This safety factor was defined by accidents in which a motor vehicle struck rail equipment either at dusk, dawn, or dark, possibly because of the drivers' inability to recognize or detect the train (Carroll, et al., 1999).

Our analysis compared the change in driver decision making from 1986 to 2007 using data provided in the FRA Highway-Rail Grade Crossing Accident/Incident database. This database contains details on each accident that occurred at a highway-rail grade crossing per calendar year. The year 1986 was selected as a baseline for consistency with Raslear (1996), and it reflected a time period before any of the safety factors had been introduced. The year 2007 was used because

it was the most recent year for which data was available when this analysis was conducted. The one exception to this comparison was “improving warning device reliability”. In 1997, the FRA modified its reporting structure associated with the reporting of that safety factor to include five additional criteria; as a result, more accidents were identified that were attributed to that safety factor. Given the inconsistency in the information collected on warning device reliability between 1986 and 2007, we were concerned that comparing the two years would be misleading. We chose instead to compare the 2007 data to information collected in 1997 (the first year in which the new criteria were used).

The data fields in the FRA Grade Crossing Accident/ Incident database describing an accident can cite several safety factors contributing to an accident. For example, an accident involving a motor vehicle that was stalled on the tracks and was hit by a train would be included in the accident count for two of the safety factors listed above: “improving locomotive conspicuity through the use of alerting lights” and “improving the reliability of motor vehicles”. To avoid over-counting of accidents associated with a safety factor, we “isolated” the number of accidents associated with each safety factor using a method described by Horton, et al., (2008). Each accident was classified as attributable to one safety factor, more than one safety factor, or none of the safety factors. The estimated number of accidents for each safety factor was then calculated using a count of those accidents attributed to that one safety factor only.

We evaluated changes in driver decision making by estimating sensitivity and bias for eight grade crossing warning devices and examining shifts in d' and β across each of the warning devices for each safety factor. The eight grade crossing warning devices consisted of: 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.

There are several ways to calculate sensitivity and bias. For this analysis, 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 β 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., $1 - P(AC)$). $P(AC)$ was estimated for each grade crossing warning device as the accident rate per crossing per train per highway vehicle per minute to equalize for exposure.

We estimated $P(FS)$ as 1 minus the probability of a correct crossing, $P(CC)$. $P(CC)$ reflects the probability that a car and a train will arrive at the crossing at the same time and that one (or both) will stop. To estimate $P(CC)$, we quantified the reverse, that is, the probability that a car and a train will arrive at the grade crossing at the same time and that neither can stop. We described this estimate as the maximum accident risk at a grade crossing, $P(AC)_{max}$. Therefore, because $P(CC)$ can be calculated as 1 minus $P(AC)_{max}$, $P(FS)$ is equal to the accident risk associated with each warning device. That is:

$$P(FS) = 1 - P(CC) = 1 - [1 - P(AC)_{max}] = P(AC)_{max}$$

and

$$P(AC)_{max} = p_T \times p_H, \text{ where}$$

p_T is the probability that one or more trains would be observed at a grade crossing in a 1-minute period

p_H is the probability that one or more highway vehicles would be observed at a grade crossing in a 1-minute period

The values for p_T and p_H can be estimated using the train rate per day at a crossing and the average annual daily traffic (AADT) at a crossing, respectively.

It is worth noting that by estimating $P(FS)$ as the accident risk, we characterize “conservative” behavior differently than other SDT models. Conservative behavior generally reflects a reluctance to recognize that a signal is present (i.e., a greater tendency to say “no”), but it is based on an estimate of $P(FS)$ that reflects the operator’s inclination to say “yes” when no signal was present. In the grade crossing model we developed here, $P(FS) = P(AC)_{max}$, and thus reflects drivers’ tendency to say “no” when a train is present. As a result, in this framework, more conservative behavior reflects increased stopping behavior (i.e., a greater tendency to say “yes”).

The data needed to calculate $P(AC)$ and $P(AC)_{max}$ were provided in the FRA’s Highway-Rail Crossing Inventory. This database contains a list of all grade crossings, the warning device used at those crossings, the median number of trains per day for the crossings, and the median average annual daily traffic (AADT) at the crossings. However, because the fields for trains per day and AADT in the Highway-Rail Crossing Inventory are not updated each year, we adjusted our estimate for these values to more accurately reflect the increase in the overall number of trains and highway vehicles from 1986 to 2007. The increase in the number of trains was estimated by dividing the number of train miles traveled with the number of track miles for each year from 1986 through 2007 based on data provided by the FRA and the Association of American Railroads. The increase in the number of highway vehicles was estimated according to information on national VMT from 1986 through 2007. The number of trains per day and AADT was not changed for 1986, but the values for 2007 were proportionally increased to reflect the changes. (See Yeh, et al. for more details regarding this adjustment.)

It is important to note that this analysis was by no means a causal analysis. The accidents attributed to a safety factor were based on characteristics that were included in the Highway-Rail Grade Crossing Accident/Incident database, and it is not possible to determine through the data fields alone the degree to which one or more of these factors contributed to the accident.

Safety Factor	d'				β			
			ω ²				ω ²	
	1986	2007	Safety Factor	Warning Device	1986	2007	Safety Factor	Warning Device
CMV Driver Safety	7.23	7.52	0.24	0.53	0.562	0.008	0.3	0.68
Alerting Lights	7.23	7.54	0.25	0.54	0.507	0.006	0.34	0.63
Reflectorization	7.47	7.66	0	0.53	0.188	0.003	0.29	0.64
Sight Lines	7.58	7.85	0.19	0	0.142	0.002	0.28	0.63
Warning Reliability	7.61	7.74	0	0.78	0.001	0.0000916	0.19	0.68
Average	7.424	7.662	0.136	0.476	0.28	0.0038183	0.28	0.652

Table 3. Statistical summary of safety factors

RESULTS

For each safety factor, we calculated and compared estimates of d' and the β using a 2 (year) x 8 (warning device) ANOVA. Because β is non-linear, the estimates for β were analyzed using the natural logarithm. Often, the comparisons identified grade crossing warning devices with similar estimates for d' and ln β (i.e., where there was no significant difference in drivers' response to the warning devices). Thus, in many cases, the results do not lend themselves towards identifying a warning device that was "most" or "least" sensitive or "most" or "least" likely to encourage drivers to stop. The two-way ANOVAs that were performed to examine the statistical reliability of the changes in d' and ln β accruing to the introduction of the safety factor and of the grade crossing device type (collectively called grade crossing safety) also reported the proportion of variance accounted for (ω²) by the safety factor and the warning device type. **Table 3** shows the estimations for d' and β averaged across warning device and summarized for each safety factor between 1986, when the safety factors were not in effect, and 2007, when they were in use.

Overall, mean d' increased by 3.2% as a result of introducing the safety factors, while mean β decreased by 165%. That is, drivers became both more sensitive and more conservative (likely to stop) over the 21-year time period. Both changes are statistically reliable (for d', t(4) = 7.04, p < 0.01; for ln β, t(4) = 10.27, p < 0.01). Drivers demonstrated higher sensitivity at grade crossings protected by passive warning devices than at crossings protected by active warning devices but drivers behaved more conservatively at grade crossings protected by active warning devices than at crossings protected by passive warning devices.

The data generally showed statistically significant improvements in d' and β attributable to both the warning device type and the introduction of the safety factor. An examination of ω² indicates that warning devices have the most impact on grade crossing safety. A visual comparison of the ω² values for d' and ln β associated with safety factors and device type in **Table 3** highlights that the safety benefits provided by warning devices are due to their ability to encourage drivers' to stop at grade crossings. Regulations to

improve CMV driver safety, the use of alerting lights, sight lines, and reflectors were generally equally effective in improving grade crossing safety. Warning device reliability (which pertained to active warning devices only) contributed to improvements in grade crossing safety, but the effects were more muted.

While the analyses suggests that ln β played a much larger role in enhancing safety through the introduction of the safety factors, it is difficult to directly compare d' and ln β because the estimates of d' range from a minimum of 0 to infinity (i.e., 0 ≤ d' ≤ ∞), but the estimates of ln β can range from negative infinity to positive infinity (i.e., -∞ ≤ ln β ≤ ∞). The use of ω² to indirectly compare d' and ln β can avoid this problem because ω² indicates the strength of the association between a dependent variable (d' or ln β) and an independent variable (a safety factor) in a unit-less metric. We conducted a 2 x 2 non-parametric ANOVA on the ω² data in **Table 3** to determine if the differences in strength of association between d' and ln β, between safety factors and device types, and their interaction were reliable. The mean ω² for d' is 0.306, while for ln β it is 0.466. The difference in ω² between d' and ln β is statistically reliable (Wilcoxon Signed Rank Test, p = 0.0313), indicating that the association between bias and grade crossing safety was 50% greater than that for d'. In other words, the overall improvement in safety was due to both increases in ability to detect a train and a bias to stop, but a bias to stop was 50% more important.

The mean ω² is 0.208 for safety factor and 0.564 for device type. The difference in ω² is statistically reliable (Wilcoxon Signed Rank Test, p = 0.0313), indicating that the association between device type and grade crossing safety was 1.7 times greater than that for safety factor. That is, the overall improvement in safety was due to both grade crossing devices and safety factors, but grade crossing devices were nearly twice as effective.

The findings from our analysis differ from those reported by Horton, et al. (2008) who evaluated the impact of these safety factors as a function of the change in the number of accidents. **Table 4** shows a comparison of the estimates by Horton et al. and the mean ω² values calculated in the analysis reported here (across SDT metric).

As **Table 4** shows, Horton, et al. found that between 1994 to 2003, improving CMV driver safety accounted for the largest benefits in grade crossing safety, followed by increasing locomotive conspicuity through the use of alerting lights. Clearing sight lines and improving warning device reliability provided more moderate benefits. The use of reflectors was not included in their analysis. Examination of the mean ω^2 values for each safety factor show a different pattern across a larger timeframe, however, and indicate that CMV driver safety, alerting lights and sight lines were nearly equivalent in their effect on grade crossing safety. Warning reliability, which only applies to active warning devices and only for the period 1997 to 2007, was much less effective than the other factors in terms of the ω^2 metric, but more effective than indicated by the Horton et al. study. The differences in findings between these two studies therefore raise questions concerning the approach to analysis.

Safety Factor	Accident Reduction (Horton, et al.)	Mean ω^2 x 100
CMV Driver Safety	34.6%	27%
Alerting Lights	13.6%	29.5%
Sight Lines	3.6%	23.5%
Warning Reliability	3.1%	9.5%

Table 4. Comparison of safety factors between Horton et al. and this analysis

DISCUSSION

This analysis comparing ω^2 revealed some important facts about improving safety at grade crossings. Grade crossing devices, in particular active warning devices, are the most important safety tool because they increase the bias to stop. This bias to stop is more important than improving the driver's ability to detect the train, and even measures to improve train detection improve driver's bias to stop.

Our data showed greater contributions for each of the safety factors examined than that reported by Horton, et al. (2008). However, our approaches to analysis and the timeframe we considered in our analyses differed. Horton, et al. calculated the percent reduction by determining the percent change in the number of incidents attributable to the safety countermeasure for the ten year time-period from 1994 to 2003 (i.e., the change in the number of incidents for a safety factor from 1994 to 2003 divided by the total number of grade crossing incidents from 1994 to 2003). Rather than use accident frequency as in Horton, et al., our SDT model estimated metrics describing driver sensitivity and bias, which were derived using a Poisson distribution of the accident rate and examined data for a 21-year time period from 1986 to 2007. The Poisson distribution (as we used here) is often used to model accident frequencies, because accidents are rare events. Consequently, examining accident rate and/or changes

in accident frequency alone may not be a good way to assess safety or determine the significance of the change in the accident rate.

The findings from this analysis suggest that the greatest benefits to improving grade crossing safety may be obtained for those countermeasures that encourage drivers to stop at the crossing (e.g., legislation and education). Based on the findings, we believe that the SDT model described here provides a framework that can be applied to understand the impact of other countermeasures for improving grade crossing safety and is more descriptive than that offered by more "traditional" analyses (e.g., accident analyses). One of the key aspects of the framework is the consideration of accident frequency with respect to human behavioral metrics that influence driver decision making. In fact, the analysis shown here suggests that the examination of accident frequency alone is misleading and may minimize the impact of other important safety factors. Although the basic signal detection model is descriptive in nature, it can be refined in conjunction with field studies or laboratory experiments to provide a better understanding of driver behavior.

ACKNOWLEDGEMENTS

This research was conducted with funding from the FRA's Office of Research and Development while the first author was a staff member at the Volpe National Transportation Systems Center. 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

- Carroll, A., Multer, J., Williams, D., and Yaffee, M.A. (1999). *Freight Car ReflectORIZATION* (Report No. DOT/FRA/ORD-98/11). Cambridge, MA: US Department of Transportation, Research and Innovative Technology Administration., Volpe Center.
- Horton, S., Carroll, A., Chaudhary, M., Ngamdung, T., Mozentner, J., and Skinner, D. (2008). *Success Factors in the Reduction of Highway-Rail Grade Crossing Incidents from 1994 to 2003*. Cambridge, MA: US Department of Transportation, Research and Innovative Technology Administration., Volpe Center.
- 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. Helsner (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., Multer, J., and Raslear, T. (2009). An application of signal detection theory for understanding driver behavior at highway-rail grade crossings. *Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.
- Yeh, M., Raslear, T., and Multer, J. (in preparation). *Understanding Driver Behavior at Grade Crossings through Signal Detection Theory*.