THE UNIVERSITY OF CALGARY

The Effect of In-vehicle Railway Warning Reliability on Trust and Performance

by

Jasdeep S. Chugh

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF PSYCHOLOGY

CALGARY, ALBERTA

JANUARY, 1999

0-612-38574-4

Abstract

Driver behaviour at railway crossings has tragic, fatal consequences for vehicle occupants. Accidents take place because drivers, at times, are not aware that trains are coming. New technology can detect approaching trains and send warning signals to oncoming vehicles. These In-vehicle Railway Warnings (IRWs) are redundant with those in the external environment. Two in-vehicle warning issues were studied. Study 1, asked 20 drivers, when, during approaches to railway crossings, IRWs should activate. Drivers consistently chose warnings to come on 10 s before vehicles reach crossings. Study 2, with 36 drivers using a low-fidelity simulation, tested the effects of IRW reliability on motorist performance and trust. Results showed that drivers tended to initiate braking later after false alarms and earlier after missed signals. Decreases in reliability reduced drivers' trust in and ratings of dependability of the IRW system. Design of IRWs must ensure that warnings to drivers are highly reliable.

Acknowledgments

I thank the Almighty God for all the near misses that didn't result in accidents. I would like to thank my family for their support during my pursuit of a master's degree. I thank Dr. Jeff Caird for supervising my master's education. The University Research Grant Committee graciously funded part of this thesis. I thank Brad Johnson for the computer programming, and Ivan Hoza and Weiqun Qiang for building the brake/accelerator simulation. Com-Media, at University of Calgary, provided the Hi-8 Camera for summer videotaping and the TV/VCR for Study 1. The computers for digitizing and making of the movies, and the projector for Study 2 were also provided by Com-Media. Fred Houle, Blair Forrester and Mike Mattson at Com-Media were of immense help. I would also like to thank Dr. Bob Dewar, Dr. Chip Scialfa and Dr. Saul Greenberg for being part of this thesis committee. I thank Sunny Thinda for running the participants through Study 1, and Philip Hove for partnership in the pilot project for Study 1. Cam Nelson, Traffic Analyst at the Calgary Police Service, provided the accident case studies.

Table of Contents

List of Tables

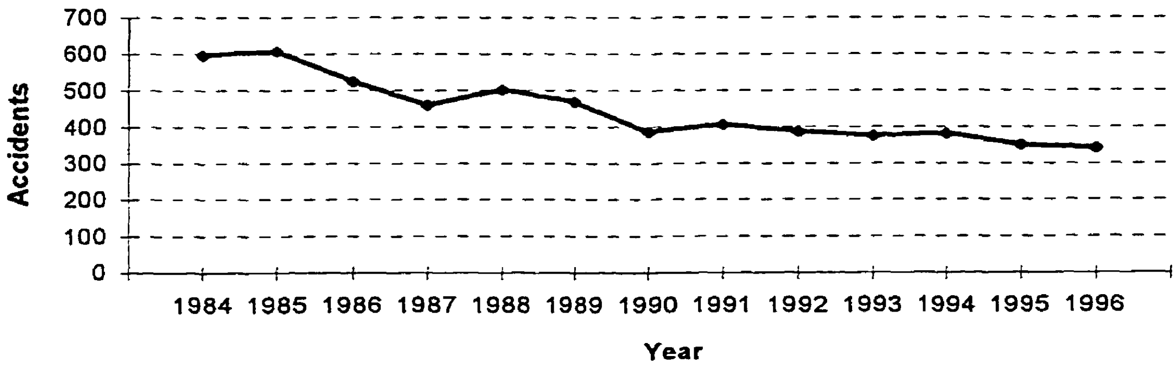
List of Figures

List of Appendices

Introduction

Trains have been transporting goods and persons for more than one and one-half centuries. The share of real estate between trains and other modes of transportation (e.g., automobiles) has resulted in countless fatalities and injuries. From 1984 to 1996, there has been a steady decline in vehicle accidents at Canadian railway crossings. Since 1990, the decrease in accidents has leveled to between 350 and 400 per year (see Figure 1) (Transport Canada [TC], 1998). Similar data collected from the US Fatal Accident Reporting System (FARS) database shows a decreasing trend in fatal accidents since 1972. However, the number of fatalities have remained somewhat constant at 400 to 500 since 1982 (see Figure 2) (Klein, Morgan & Weiner, 1994). Note that Figure 1 describes all accidents while Figure 2 describes only those accidents that resulted in fatalities.

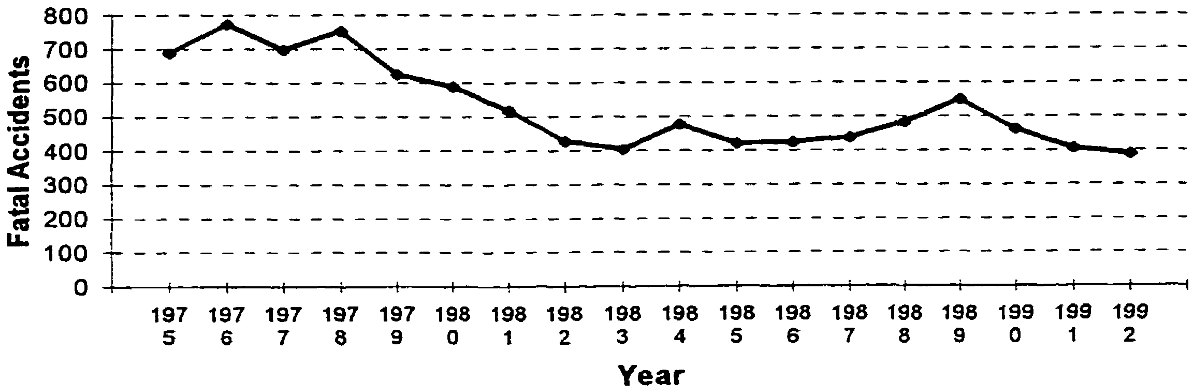
Figure 1. Accidents at Canadian railway crossings from 1984 to 1996



In the US, younger drivers (aged 16-24) are most likely to be involved in fatal accidents (see Figure 3) (Klein et al, 1994). In Canada, drivers in the 16-19 age category and those above 65 years were involved in a higher proportion of fatal accidents at crossings (TC, 1996). From 1972 to 1992, males were involved in 77% of the crossing accidents fatalities, whereas females accounted for the remaining 23% (Klein et al., 1994).

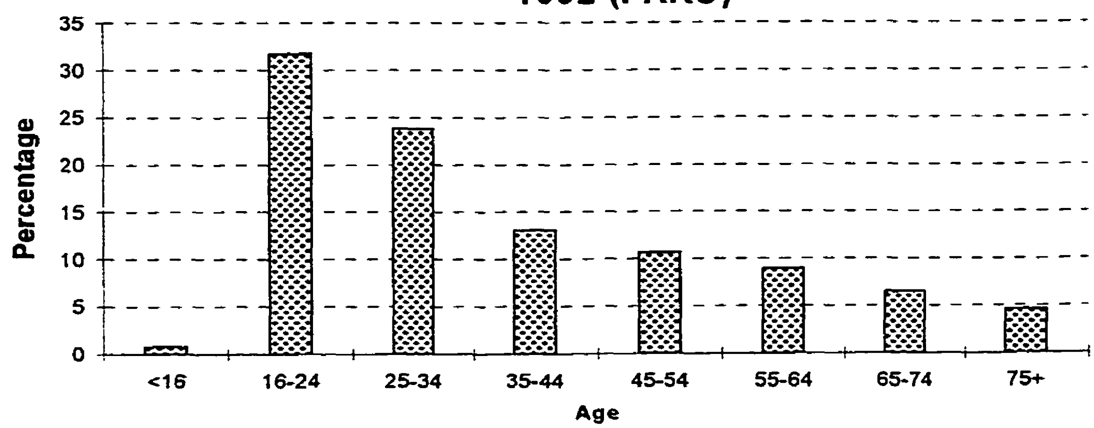
While epidemiological accident data suggests that accident rates for younger female drivers (18-23 years) have steadily increased (Cerrelli, 1994), younger males are

Figure 2. **Fatal accidents at railway crossings from 1975 to 1992 (FARS)**



still the riskiest of drivers, as they follow other drivers too closely and speed excessively (Rabinovich, 1996). As more younger drivers are involved in accidents at railway crossings, this thesis focuses on this demographic segment.

Figure 3. **Fatal accidents by driver age from 1975 to 1992 (FARS)**



In Canada, 342 crossing accidents took place in 1996 which resulted in 47 fatalities and 68 injuries (Transportation Safety Board of Canada [TSB], 1997b). Thirty-eight of the fatalities and 63 of the injuries were to motor vehicle occupants. These statistics reveal that motor vehicle occupants are at the greatest risk of being in a train-related accident and experiencing injuries, either fatal or non-fatal, at railway crossings

(TSB, 1997c). TSB (1997a) notes that, "[m]otor vehicle driver behaviour plays a major role in most crossing accidents" (p. 3). A motorist is 30 times more likely to be fatally injured in a collision involving a train than in a crash involving another motor vehicle (Operation Lifesaver, 1997). This statistic is of grave importance, given the relatively low number of interactions between motorists and trains compared to interactions between motorists.

Most crossings have crossbuck signs (passive control), and may also have active devices (e.g., warning bells, lights and gates). The crossbuck is categorized as a passive warning because it identifies a set of tracks intersecting the roadway, and not necessarily that a train is coming. As lights and bells are activated when trains near crossings they are categorized as active warnings. In 1996, 52% of crossing accidents and 67% of fatalities took place at all types of active crossings (TSB, 1997a-d). These active crossings comprise 31% of Canadian public crossings (see Table 1). Seven percent of all active crossings had gates, in addition to bells and lights. Based on 1996 statistics, fatalities and accidents are much more prevalent at active crossings, both with and without gates, than at passive crossings. The difference between fatal accidents at active crossings and at passive crossings may be because of the increased likelihood of train-motor vehicle interaction at active crossings. Other factors, such as restricted sight lines may also make them more dangerous.

Table 1

Canadian public railway crossing accidents for type of crossing in 1996 (TSB, 1997d)

Type of crossing	**Number of public crossings (%)**	**Accidents (%)**	**Fatalities (%)**	**Injuries (%)**
Passive (crossbuck only)	15,686 (69%)	139 (48%)	14 (33%)	31 (50%)
Active (bells and lights)	5,642 (25%)	120 (42%)	18 (43%)	23 (37%)
Active (bells, lights & gates)	1,498 (7%)	29 (10%)	10 (24%)	8 (13%)
Total	22,850 (100%)	288 (100%)	42 (100%)	62 (100%)

In the US in 1995, there were 3,972 motor vehicle - train accidents/incidents (FRA, 1996) resulting in 455 fatalities and 1,696 non-fatal injuries. In 2,947 or 74% of the accidents/incidents, trains struck the vehicles. Three hundred and ninety-one (86%) of the fatalities and 1,197 (71%) of the injuries were due to trains striking motor vehicles. This finding suggests that most fatalities/injuries are the result of drivers violating trains right-of-way, before trains arrive at crossings. The occurrence of these accidents also suggests that such violations take place only a few seconds prior to the arrival of trains.

Of the 2,947 accidents where trains struck motorist vehicles, 62% or 1,828 occurred during daylight conditions, with 32% or 932 during nighttime (FRA, 1996). The remaining 6% took place during dawn and dusk times. The highest number of accidents/incidents took place from 11 A.M. to 6 P.M., during clear visibility conditions with no obstructions of the drivers' views of the crossings. There are also more people driving during daylight conditions, resulting in increased exposure to trains. Of all accidents/incidents, January and December had the highest number, followed by October, August and February. As the number of trains per day increased, along with an increase in motorist traffic flow, so did the number of accidents. Higher than average number of accidents took place in urban and commercial settings than in rural and residential areas.

Crossing accidents occurred over a wide range of train speeds, with more casualties taking place at higher train speeds (FRA, 1996). Motor vehicles were traveling at a range of speeds, from stopped or stalled on the tracks to as high as 80 kph, with more accidents taking place at slower motorist speeds. Most of the motor vehicles involved in such collisions were classified as either automobiles or trucks (3,540 or 89%). Three school buses and 412 truck-trailers were also involved in crossing accidents.

In summary, the US (1995) motor vehicle-train accident statistics indicate that trains strike vehicles in a variety of environmental conditions. These accidents take place at crossing locations, that while heavily traveled, are also heavily equipped with both passive and active traffic control devices. Most of Canadian train-involved fatality accidents are with motor vehicles, with most of the fatalities being those of motor vehicle occupants. Drivers are usually cited for committing violations that resulted in the accidents. Younger drivers are least likely to comply with warning devices at railway crossings.

Along with Canadian and US train-vehicle accident databases, 11 cases of train-vehicle accidents within Calgary city limits were also reviewed. These accidents occurred between 1993 and 1996. These 11 were not the only train-vehicle accidents that took place during these years, but were those for which non-sensitive information was available from the City of Calgary Traffic Section (Nelson, 1998). Thus, they should be interpreted with caution. The 11 drivers ranged in age from 22 to 78 years. In 4 of the 11 cases, drivers had driven into restricted areas, and had either struck parked trains or were struck by moving trains. The remaining 7 accidents took place at railway crossings between 11 A.M. and 6 P.M. In these cases, the drivers violated the right-of-way of moving trains. One involved an American tourist unfamiliar with the downtown area (specifically one-way streets), and another was a police officer driving to an emergency situation. The other 5 drivers were cited for failing to stop at a red light, failing to stop at a crossing signal and/or driving without due care or attention. Two of these 5 were older drivers (ages 73 and 78 years). Some similarities (e.g., time of day, presence of active devices, trains striking vehicles) in spite of the low number of cases, are apparent between local and national statistics.

Driver Behaviour at Railway Crossings

In 2,212 (56%) of the 1995 US accidents, drivers were cited for not stopping before crossing railway tracks. Another 493 (12%) of the accidents occurred when motorists drove around or through lowered gates, and in 189 (5%) drivers stopped and then attempted to proceed through the tracks (FRA, 1996). The rest, 1078 (27%), were classified as *other* or *unknown*. These statistics suggest that, for most accidents, drivers either made incorrect decisions to cross the tracks before trains arrived, drivers were unable to stop their vehicles before reaching crossing locations, or drivers did not know that trains were approaching the crossings. The motorists that drove around the gates clearly violated the right-of-way of trains.

Canadian statistics, from 1991 to 1994, reveal similar accident causes. Among drivers involved in fatal crashes at railway crossings, 44% had disobeyed traffic control signals, 27% had failed to yield right-of-way, and another 5% were driving too fast (TC, 1996). Clearly, driver behaviour plays a major role in crossing accidents (Leibowitz, 1985; TSB, 1997a). While these aggregate statistics show that drivers were at fault or were cited for violating right-of-way of trains, it seems from the police reports that for all Calgary cases, the drivers did not know that trains were coming. These accidents occurred even though six of the seven crossings had active warnings (i.e., traffic lights at red, flashing lights, ringing bells, lowered gates) in place and in proper working order.

A number of researchers have described driver behaviour at railway crossings (Åberg, 1988; Abraham, Datta & Datta, 1998; Knoblauch et al., 1982; Leibowitz, 1985; Lerner, Ratte & Walker, 1990). Åberg (1988) made approximately 2,000 observations of driver behaviour at approaches to 16 different crossings. He found that head movements to look for trains from the left and the right decreased as visibility at crossings was reduced. In other words, if there was an object (e.g., building, vegetation) that restricted a driver's visibility to an approaching train's path then the driver did not look in that direction. Similarly, when the angle of the crossing was high (i.e., close to 180°), then drivers did not make efforts to turn and look for approaching trains.

While Åberg (1988) acknowledges that failing to look for trains does not necessarily mean warnings and signals were not noticed, he states that other researchers (e.g., Thorson, 1976; Wigglesworth, 1979) have analyzed accident reports showing that a "typical victim in a train-vehicle accident had been observed to drive steadily, without any head movements, straight in front of the train, in spite of activated flashing lights and alarm bells" (p. 64, Åberg, 1988). During these instances, drivers may be distracted by internal or external factors (Knoblauch, Hucke & Berg, 1982). These observations apply to both actively and passively controlled crossings. Accident data also show that even though some crossings are equipped with active warning devices, the same number of accidents take place there as they do at passively controlled crossings.

Even when drivers are observant, and notice warnings and signals, they do not fully comprehend the factors that contribute to accidents (Leibowitz, 1985; Lerner et al., 1990). For example, some drivers believe that train operators can slow down and stop trains if they see that vehicles are on the tracks (Richards & Heathington, 1988). What drivers do not know is that it takes large trains traveling at 45 mph approximately one mile to stop (Operation Lifesaver, 1997). A train operator doesn't know if a driver is going to try to cross the tracks before the train gets there, and thus, whether to brake or not. When a motorist tries to beat a train, then, at times, he/she is struck. While violations of traffic rules at railway crossings are infractions, Abraham et al. (1998) did not encounter any law enforcement activities during an 18 month study.

Leibowitz (1985) examines a number of reasons why drivers attempt to cross tracks in front of approaching trains. There is a degree of uncertainty of when a train will actually arrive at the crossing, relative to when the bells and/or light warnings are activated. This is because drivers are unable to accurately predict speeds of trains (Meeker, Fox & Webber, 1997). There is also uncertainty about how long the wait will be before a train completes its crossing. Long delays have been known to occur (Abraham et al., 1998), especially at crossings that are located near rail yards (Leibowitz, 1985). In not wanting to wait for an unknown length of time, some drivers decide that it is better to try to cross before the trains arrive at the crossings. Thus, the uncertainty of delay coupled with the belief that train operators can stop trains gives reason to some drivers to violate right-of-way of trains.

Driver behaviour at railway crossings is not all bad. Drivers do approach crossings more slowly at nighttime (when visibility is poorer) than during the day (Ward & Wilde, 1995a). Drivers do drive slower near crossings when sight lines are restricted, but tend to increase speed when sight lines improve (Ward & Wilde, 1996). Ward and Wilde (1995b) also showed that drivers can be cued by external signs to slow their speed when approaching railway crossings. The researchers had actually placed signs that asked drivers to stop at crossings. As no trains were present in the vicinity and as no justification was provided to drivers for the request, they did not stop. They did, however, slow down.

Continual compliance to signs requires justification for the requests. For example, a sign saying SLOW YOUR SPEED AS A TRAIN MAY BE COMING tells drivers what to do and why they should do it.

Review of driver behaviour has shown that a number of factors contribute to accidents at railway crossings. One of them is lack of driver awareness about approaching trains. If drivers are not aware that trains are coming, then there is the need for more redundant warnings. A new resource that can be used to warn drivers in a timely manner about approaching trains is an In-vehicle Railway Warning (IRW) system. In an IRW system, wireless signals are transmitted from crossings to approaching vehicles, which then display an IRW to drivers. New technology within vehicles allows display of warning signs on the windshield (Killick et al., 1997), similar to head-up displays (HUD) (Caird & Chugh, 1997). Thus, an IRW can appear in a HUD.

The goal of these IRWs is to reduce the number of accidents. IRW systems can be tailored to both motorists' capabilities and limitations, and the operating environment (Tufano, Knee & Spelt, 1996) in a way that traditional signing cannot. For example, display location and duration of information presentation may be tailored to suit specific individual needs. What is to be displayed, and how and when to display it can be designed to fit a myriad of purposes. Given the frequency of collisions between trains and motorists with existing warnings, newer technology use at crossings could result in benefits of increased safety and reduced collisions (Richards & Bartoskewitz, 1995).

This thesis focused on two important issues with IRWs. The first study addressed where in approach of crossings IRW should be activated. While the goal of an IRW system is noble, the possibility of reduced reliability may diminish proposed benefits. Thus, the second study looked at the effects of reduced reliability on motorist performance and trust.

Study 1

Current Warning Devices

There are a number of traffic control devices (TCDs) in place at railway crossings. A crossbuck railway crossing sign is installed at all crossings. The crossbuck sign is a post-mounted warning, and placement requirements for the crossbuck are specified in the *Canadian Road/Railway Grade Crossing Manual* (Road Safety Directorate [RSD], 1995). Other signs depicting the number of tracks and advisory approach speeds can also complement crossbuck signs, but are not required at all crossings. At some crossings, crossbuck signs are supplemented with flashing lights and bells (RSD, 1995). These devices are turned on when trains are approaching crossings and remain on until trains have passed through. Some crossings also have gates that lower to physically block motorists or pedestrians from crossing the tracks.

Another passive warning is the Railway Advance Warning Sign (R-AWS). The R-AWS is a pictorial representation of a roadway intersecting a railway (code WC-4). These signs are placed 50 to 150 m from crossings (Section A1.15, CUTCD, 1976). R-AWS are installed in advance of crossings whose crossbuck signs are not clearly visible within normal stopping sight distances (CUTCD, 1976). Some crossings may not require R-AWS if they are in commercial sectors of urban areas and are equipped with automatic warning devices (i.e., bells, lights, gates). Supplemental to R-AWS, often there are large X-pavement markings on the surface of roadways in advance of crossings. Similar to crossbucks, R-AWS and X-pavement markings identify upcoming crossings.

There can also be Automated Advance Warning Signs (AAWS) placed in advance of railway crossings. Similar to bells and lights, AAWS are activated by approaches of trains. AAWS are installed at crossings that have poor sight distances for other active warnings (i.e., bells, lights and gates). AAWS are also installed in advance of crossings which are in areas susceptible to frequent fog, and where road approaches steeply descend towards crossings (CUTCD, 1976). In some locations, the activation of AAWS may be initiated before the bells and lights at the crossings are turned on to provide motor vehicle drivers with more time to slow down. AAWS may also remain on after the crossing warnings have been terminated to provide vehicles stopped at the crossings time to re-accelerate before following motorists arrive at the crossings.

The *Alberta Basic License Driver's Handbook* provides learning drivers with information about signs placed at railway crossings (Alberta Transportation and Utilities [ATU], 1997a). Crossbucks mean that drivers have to yield to all trains. R-AWS tells motorists to look, listen and slow down as they may have to stop (ATU, 1997a). The R-AWS is in a set of advance warning signs that include stop ahead, traffic signals ahead and pedestrian crossing ahead. The handbook guides drivers to be especially alert for R-AWS during poor weather and nighttime conditions. When lights are flashing and bells are ringing at crossings, then drivers must always stop, and only proceed (a) when the lights/bells have stopped, (b) when trains have passed through or (c) when drivers feel they can proceed safely. If there are two or more railway tracks intersecting the roadways, then drivers should make sure all tracks are clear before crossing. Drivers are directed to not race trains to crossings, and to always stop when approaching trains are visible and within 500 m of crossings. A supplement handbook for professional drivers (e.g., truckers and bus drivers) guides them to slow to speeds that will let them stop safely when approaching any railway crossing (ATU, 1997b). Bus drivers must stop at all railway crossings that are not equipped with active warnings (i.e., bells, lights, gates).

Positive Guidance

The placement of TCDs is advised by Positive Guidance, which is the theoretical basis of highway design (Alexander & Lunenfeld, 1975). It is based on the premise that drivers can be provided with timely information about upcoming hazards, through formal and informal cues, allowing them to operate motor vehicles safely and error-free. Hazards can be from both fixed and moving objects. Formal cues include warning and regulatory signs, delineations and pavement markings. Informal cues include road geometry, other vehicles, buildings and vegetation along the roadway. There are five information handling zones in which formal cues about hazards are placed (Alexander & Lunenfeld, 1984, pp. 363-364). The lengths of each of these zones, in terms of distance and time to pass through, are tailored to individual hazard locations using engineering principles (e.g., vehicle speed). The five zones are:

a) The *Advance Zone*. The advance zone is the area in which an upcoming hazard does not pose a threat, but in which advance warning(s) to inform drivers can be effective.
b) The *Approach Zone*. The approach zone is the area in which an upcoming hazard can be perceived, at a point termed as the *decision sight distance*, and where warning(s) pertaining to the hazard can be placed.
c) The *Non-recovery Zone*. The non-recovery zone is the area in which there is insufficient time to avoid an interaction with a hazard, and where information regarding the hazard should not be placed to avoid overloading drivers.
d) The *Hazard Zone*. The hazard zone is the area in which the interaction between a hazard and motor vehicles takes place.
e) The *Downstream Zone*. The downstream zone is the area beyond a hazard in which drivers can relax their vigilance and resume their course.

The five zones also exist in approaches to railway crossings (Lerner et al., 1990). Placement of warnings (e.g., R-AWS) should be in the approach zone, and the range of 50 to 150 m corresponds to this zone (Section A1.15, CUTCD, 1976). If a R-AWS is placed at 100 m prior to a crossing, for example, it is seen by drivers before they reach that location. Earlier perception of the R-AWS expands the 50 to 150 m range. Crossbuck signs, lights, bells and gates are placed in hazard zones, and they identify the location at which hazards (i.e., trains) cross roadways.

Lerner et al. (1990) suggest that measurements of the five information zones used by engineers may not necessarily incorporate drivers' needs and behaviours. Transition from one zone (e.g., approach) to the next (e.g., non-recovery), for example, may be based on inflexible estimates of driver reaction time, vehicle braking distance, visibility conditions, etc. Therefore, "[i]n reality, any particular combination of driver/vehicle/trip/environment may define different transition points between zones" (Lerner et al., 1990, p. 2-3). Subsequently, decision sight distance criteria, important for placement of warnings, in the approach zones may also be susceptible to variation dependent on a variety of factors (e.g., roadway geometry, driver characteristics).

Decision Sight Distance

For drivers to be able to respond to moving hazards effectively, warning information in the approach zones should be at decision sight distance points. Decision sight distance is the distance required by drivers to perform five information handling operations. The five operations are: searching for a hazard, detecting it, recognizing it as a hazard, selecting appropriate speed and path, and following through a chosen maneuver safely and efficiently. Search time for a hazard is included in decision sight distance especially for those cases when the hazard is not coming towards drivers from the front, but from the side (Alexander, 1991).

Olson (1996) provided the recommended time values, in seconds, for decision sight distance criteria. The detection and recognition of a hazard on a roadway requires 1.5 to 3 s. A decision to select an appropriate path or maneuver varies from 4.2 to 7.0 s. The time to make the maneuver (e.g., a lane change) is 4.5 s. The overall decision sight distance is then in the range of 10.2 to 14 s. Olson (1996) did not include search time in decision sight distance criteria. The inclusion of search time would increase the 10.2 to 14 s range (Alexander, 1991). The distance traveled by a vehicle in 10.2 and 14 s is dependent on vehicle speed. At 40 kph, distance traveled in 10.2 and 14 s is 113 and 155 m, respectively. The distances rise to 297 and 389 m when vehicle speed increases to 100 kph. If drivers have more than one option on appropriate paths or maneuvers, then decision time rises (Olson, 1996). For example, if a driver notices that a construction truck is coming out into his/her lane from a side road, then the driver has two options. Either the driver can make a lane change or he/she can brake to avoid hitting the truck. Deciding between options increases decision time. Thus, the placement of TCDs, for roadways in general and at railway crossings in particular, should consider decision sight distances (Alexander & Lunenfeld, 1975), and include time required to search for oncoming trains from both the left and the right sides of drivers (Alexander, 1991).

Olson (1996) specifies the lane change as a maneuver that can be incorporated into decision sight distance calculations. Other maneuvers, like stopping, are not mentioned nor are they assigned time values. There is a separate engineering principle of *stopping* *sight distance*, which does provide time and distance values for stopping a vehicle that is traveling at different speeds. These calculations are, however, based on a driver's pre-knowledge that the response to be made is the brake press. If a driver does not have *a priori* knowledge that a brake response is to be made then decision sight distance criteria is applicable, even if the appropriate response is to press the brake.

Detection/Comprehension of Warnings

When drivers come towards railway crossings, then they initially need to detect the presence of railway crossings in which the train hazards pass (RSD, 1995). Railway crossings are identified by R-AWS in the approach zones. Factors that can influence detection of R-AWS include search and attention conspicuity. Search conspicuity is the time it takes to detect an object when it is searched for, while attention conspicuity is the time it takes an object to draw attention when it is not explicitly searched for (Lerner et al., 1990). The conspicuity of a warning (e.g., R-AWS) is a function of various characteristics of the warning embedded in the surrounding environment (Lerner et al., 1990). If a driver is not actively searching for a R-AWS and if the R-AWS is embedded in a high clutter environment, then the possibility of detection decreases (Lerner et al., 1990).

According to Lerner et al. (1990), drivers also need to identify if the railway crossings are passively or actively protected. If a crossing is passively controlled (i.e., is equipped only with a crossbuck sign), then different response strategies are elicited than when active warnings are present (e.g., warning bells and lights). If crossings are passively protected then drivers should search for approaching trains, while if they are actively controlled then drivers should look for the activation of the bells, lights and/or gates. Canadian R-AWS do not identify type of crossing (either passive or active). However, a R-AWS that depicts type of crossing is used in Australia (Standards Australia, 1991).

Drivers also need to detect the exact location of crossings (Lerner et al., 1990; RSD,1995). Through repeated exposure to railway crossings and R-AWS, drivers may come to know that detection of R-AWS means that crossings are 50 to 150 m away. The final detection by a driver is of an approaching train. When a crossing is actively protected, then the active device(s) will inform a driver that a train is present. The driver still, however, needs to identify the direction the train is traveling.

Following detection is comprehension of warning message. Comprehension of the Canadian crossbuck sign was found to be quite high in paper and pencil tests (Ells et al., 1980). While comprehension is high, Hawkins (1994) reports that some drivers (54% in their survey) do not know that R-AWS are placed ahead of crossings rather than at the crossings. Drivers are, however, knowledgeable that crossbucks and other advance warning signs inform them that there are railway crossings up ahead.

While detection of warnings and comprehension of messages are important precursors to selection of maneuvers to avoid hazards, observations of driver behaviour (see Åberg, 1988) show that, at times, drivers do not notice passive and/or active warnings, encroach onto crossings and are struck by trains. The lack of awareness suggests that detection of crossings, type of control, exact crossing locations or trains was not made. The fact that a large number of accidents, 52%, take place at crossings that had some form of active device in place (TSB, 1997d), lends support for these observations. Thus, more warnings (e.g., IRW), redundant with existing warnings, are needed.

Previous Research with In-vehicle Signs

One of the few tests with in-vehicle signs was the use of advance cueing for signalized left turn intersections (Staplin & Fisk, 1991). Advance cueing, at locations prior to intersections, on left turn rules, improved decision accuracy and latency (i.e., time to response) for both young and older participants. While participant responses to the decision task improved, an increase in tracking task errors was also found. The distraction of drivers by in-vehicle tasks (e.g., responding to a display) is an important design and research issue with in-vehicle signs. While driving, responding to an in-vehicle display about left turn rules, for example, is secondary to the primary tracking task. Staplin and Fisk (1991) suggested that to avoid complex provision of advance cueing information, and therefore to avoid distraction from the main driving task, in-vehicle information needs to be sufficiently conspicuous and legible for it to be perceived and understood at a glance.

An IRW's ability to attract attention, shift it to the primary hazard (i.e., a train), and distract the driver from performing secondary tasks (e.g., talking on the cellular telephone) will be a benefit. Drivers should, however, not be distracted from performing other primary tasks (e.g., tracking). If IRWs are only activated when trains are coming, then, through repeated exposure, drivers will form expectancies to become prepared to stop for approaching trains every time such warnings are activated. Such expectancies can only be established if and when the IRW are conspicuous and unambiguous enough so that they are easily detected and recognized when they appear. Study 1 tested for when IRWs should activate in approaches to railway crossings.

Present Study

The purpose of Study 1 was to test if driver gender, seasonal conditions, and presence of external warnings effected responses of when IRW should appear. Twenty participants viewed roadway approaches to 16 railway crossings (8 winter and 8 summer), and indicated, in approach streams, optimal locations for display of IRW about oncoming trains. The 3-way experimental design for Study 1 was Gender (male/female) x Season (summer/winter) x External Warnings (present/absent). Gender was chosen as a variable because males are involved in more crossing accidents than females. In winter, the driving environment is degraded as there is snow on the roadway. Therefore, Season was chosen as a variable. External Warnings was chosen as a variable because some railway crossings are equipped with existing warnings while others are not.

Hypotheses

1. Winter responses would be longer than summer responses.
2. Males would select locations closer to crossings than females.
3. Responses would be influenced by presence of external warnings.
4. Responses would reflect that IRW act as pre-cues to existing post-mounted R-AWS.

Study 1 : Method

Participants

Twenty volunteers (10 male and 10 female) from the University of Calgary participated in Study 1. Their ages ranged from 18 to 22 years ($\underline{M}$ = 20.5 years). All held valid driver licenses and had been active drivers for the last 3 years, with an average of 12,500 km driven per year. Half of the participants (5 males and 5 females) wore corrective vision lenses (i.e., glasses or contacts) during Study 1 testing. The mean corrected visual acuity was 20/16 and contrast sensitivity was normal, as tested at 36 cm. All were familiar with the operation of a video cassette recorder (VCR) and a remote control. They were paid $5 each for participating.

Data from a pilot study (Chugh & Hove, 1997) were used to calculate effect size and power to estimate the appropriate sample size required for Study 1. Effect size of gender from the pilot study was calculated ($\underline{d}$ = 1.15) with male ($\underline{M}$ = 9.9 s) and female ($\underline{M}$ = 8.4 s) mean responses and their common standard deviation, $\underline{SD}$ = 1.31. With a between group size of 10 (i.e., 10 males and 10 females in each gender group), power of .87 would be observed for a one-tailed test, and of .78 for a two-tailed test (Cohen, 1977). Thus, for Study 1, the power analysis suggested a sample size of 20 persons, 10 in each gender group.

Materials

A Hi-8 camera was used for summer (August, 1997) taping, and a VHS for winter (January, 1998) recording. The cameras were placed on a tripod between the driver and passenger seats in a 1987 Chevrolet Sprint. The video recorders were directed through the windshield at approximately the same line of sight as the driver. Vehicle speeds were kept at levels corresponding with posted speed limits and/or flow of surrounding traffic (see Table 2). Eight railway crossings were videotaped twice; once each in the summer and winter. The eight railway crossings differed in the presence of R-AWS, X-pavement markings, approach configuration, type of active devices, and presence of traffic signals. Table 2 lists these characteristics. Four of the crossings had existing advance warnings

Table 2

Rail crossing approach characteristics for eight experimental clips

Railway Type	Location	Approach Configuration	Advance Sign (#)	Pavement Markings	Control System	Traffic Signal	Season
Commercial	52 St. at 50 Ave. SE Northbound	Straight	WC-4 (1)	X - advance	X-buck Over	No	Both
Commercial	52 St. at Hubalta Rd. SE Northbound	Straight	WC-4 (1)	X - advance	X-buck	No	Both
Commercial	Glenmore Tr. at 28 St. SE Westbound	Straight	WC-4 (1)	X - advance	X-buck	No	Both
Commercial	Glenmore Tr. at 28 St. SE Eastbound	Curved	WC-4 (2)	X - advance	X-buck Gates / No-Over	No	Both
Commercial	17 Ave. at 16A St. SE Eastbound	Straight	NONE	X - advance	X-buck Gates / No-Over	No	Both
Light Rail	4 Ave. onto 10 St. SW Westbound	Straight	NONE	NONE	X-buck	Yes	Both
Light Rail	25 Ave. at McLeod Tr. SE Westbound (Erlton)	Curved	NONE	NONE	X-buck Gates	After	Both
Light Rail	14 Ave. at 19 St. NW Westbound (Lions Park)	Straight	NONE	NONE	X-buck Gates	After	Both

Note. X-buck: Crossbuck Sign; Over: Overhead lights and bells; No-Over: No overhead lights and bells.

(i.e., R-AWS and X-pavement markings) and four did not. The direction of travel identified in Table 2 for each crossing is the direction in which motorists travel. For example, 52 Street at 50 Avenue SE Northbound means that the video was recorded while driving on 52 Street SE heading north. At some crossings, there were traffic signals nearby (usually immediately after the crossings), which regulate motorist traffic flow. These traffic signals, at indeterminable times, pre-empt motorists' signals (i.e., change them from green to amber to red) in favor of approaching trains.

Killick et al. (1997) suggested that in-vehicle warnings may be useful in low visibility conditions, at busy intersections and work zones, and where fatigue may adversely affect drivers' abilities. The railway crossings in Study 1 were selected based on the recommendations of Killick et al. (1997). For example, the Glenmore Trail SE locations are crossed extensively by heavy trucks, whose drivers may experience fatigue from long hauls. The 14 Avenue at 19 Street NW location links a residential sector to a busy commercial environment. The 17 Avenue SE eastbound crossing is located in an inner city neighborhood with a predominantly elderly population. The 17 Avenue SE crossing is also susceptible to heavy traffic, late at night, after local hockey games in a nearby arena. The railway crossings were also selected based on approach configuration; such that some approaches curved, while others followed a straight road towards the crossings. All 16 video recorded approaches ended at either light rail or commercial rail crossings. The video clips ranged from 14 s to 31 s in duration, with average duration of 23 s. A Pearson correlation between summer and winter clip lengths was 0.12, and was not significant ($\underline{p}$ = .773).

Environmental conditions for summer videotaping were sunny, under a clear sky, during mid-afternoon hours, with no debris on the roadways. Winter videotaping took place during late morning/early afternoon hours, under cloudy conditions with some snow fall (<5 cm). In the winter, the roads were covered with snow and ice, and most cars were emitting steam from their mufflers which further reduced visibility. Because winter driving was on roadways covered with snow, and because driving is slower in the winter, vehicle speed during winter recording was consistently slower than the summer (see Table 3). Table 3 shows when, prior to reaching the crossings, the R-AWS and the pavement markings were crossed by the test vehicle. For example, during the summer, the R-AWS at the 52 Street at Hubalta Road SE (Northbound) crossing was passed 8 s before reaching the crossing at vehicle speed of 60kph. The same R-AWS would be passed at 10 s during the winter at vehicle speed of 50 kph.

The video clips were shown on a 69 cm (21") Sony Trinitron color television (Model: KV-27V10) and played on a Sony VCR (Model: SVO-1420). Once each clip was shown in its entirety, participants interacted with the VCR through a Sony remote control (Model: V191A) to make a response (described below).

Table 3

Vehicle speed and location of R-AWS and pavement markings

Location	Speed (kph) (Summer)	Speed (kph) (Winter)	R-AWS (#) summer / winter	X Markings summer / winter
52 St. at 50 Ave. SE Northbound	65	55	WC-4 (1) 11.75 s / 13.5 s	X - advance 11 s / 12.5 s
52 St. at Hubalta Rd. SE Northbound	60	50	WC-4 (1) 8 s / 10 s	X - advance 6 s / 8 s
Glenmore Tr. at 28 St. SE Westbound	50	45	WC-4 (1) 9 s / 10 s	X - advance 7 s / 8 s
Glenmore Tr. at 28 St. SE Eastbound	60	55	WC-4 (2) 13 s / 14 s	X - advance 11 s / 12 s
17 Ave. SE at 16A St. Eastbound	40	35	NONE	X - advance 3 s / not visible
4 Ave. onto 10 St. SW Westbound	40	35	NONE	NONE
25 Ave. at McLeod Tr. SE Westbound (Erlton)	55	45	NONE	NONE
14 Ave. at 19 St. NW Westbound (Lions Park)	40	30	NONE	NONE

Procedure

At the start of a session, participants filled out consent forms and demographics questionnaires (see Appendix A). Visual acuity and contrast sensitivity were tested using a Landolt C Near (36cm) Visual Acuity Chart and a Vistech Vision Contrast Test System (Type B) Chart, respectively. Participants were seated 70 cm from the television monitor,

Table 4

Rail crossing approach characteristics for four practice clips

Railway Type	Location	Approach Configuration	Advance Sign (#)	Pavement Markings	Control System	Traffic Signal	Season
Commercial	Barlow Tr. At 25 St. SE Southbound	Straight	WC-4 (1) 11 s	X - advance 13 s	X-buck Gates / No-Over	No	Summer
Commercial	8 St. & 9 Ave. SE Northbound	Curved	NONE	X - advance/at 5 s	X-buck Gates / Over	No	Winter
Light Rail	1 Ave. at 9A St. NW Westbound (Sunnyside)	Straight	WC-4 (1) 6 seconds	NONE	X-buck Gates	No	Winter
Light Rail	32 Ave. at 36 St. NE Westbound	Straight	NONE	NONE	X-buck Gates	Yes	Summer

Note. X-buck: Crossbuck Sign; Over: Overhead lights and bells; No-Over No overhead lights and bells.

with eye position at the center of the monitor. Seventy centimeters is the approximate distance of the windshield from drivers' eyes, and simulates the distance at which HUD warnings are displayed (Gish & Staplin, 1995).

The 20 participants were assigned to one of five orders of presentation for the 16 video clips. A description of the tasks that were to be performed was given (see Appendix B for verbal protocol). As each video clip was presented, participants were asked to imagine themselves as drivers of the vehicle in the video clips. First, a participant would watch an approach to a railway crossing in its entirety. Second, he/she would interact with the video clip using a remote control, forwarding and/or rewinding the clip selecting the location at which an IRW about an approaching train would appear. After making a response, the participant was asked to verify if the location was where he/she wanted the IRW to appear on the windshield. Each participant was given an opportunity to make another selection once for each crossing. Four example clips (two summer and two winter; see Table 4) were shown at the beginning of the experiment for practice. At the conclusion of the sessions, participants were debriefed about the study and paid.

Study 1: Results

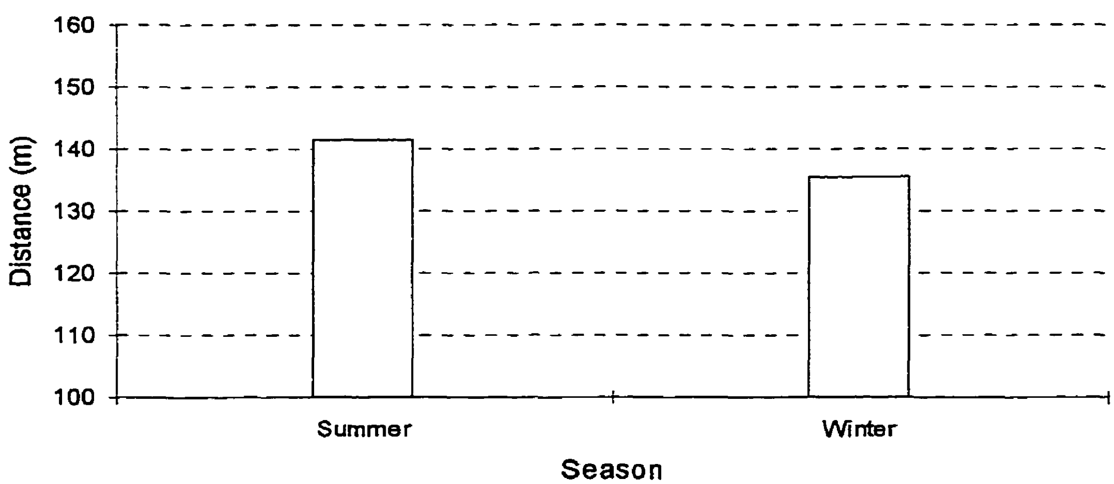
Parametric Tests

A three-way Analysis of Variance (ANOVA), of Gender (2) x Season (2) x External Warnings (2), was performed on participant responses. To avoid the confound of speed differences between summer and winter recording, the dependent measure of time (in seconds) from crossings was converted into distance (in meters). The omnibus ANOVA showed no significant three-way or two-way interactions. The main effect of Season was significant ($F(1,18) = 9.26$, $p = .007$) with more distance indicated for summer ($M = 141$ m, $SD = 24$) than winter ($M = 136$ m, $SD = 22$) conditions (see Figure 4). While a statistical difference of 5 meters was found, it was in a direction opposite than predicted by the first hypothesis. Gender was not significant as a main effect ($F(1,18) = 0.014$, $p = .9$). Thus, the second hypothesis was not supported.

The main effect of External Warnings was significant for distance ($F(1,18) = 147.87$, $p < .001$), with more distance indicated for crossings with external warnings present (M = 159 m, SD = 27) than for crossings without external warnings (M = 118 m, SD = 19). While converting time responses to distance had eliminated the confound of speed difference for Season, the conversion resulted in a confound for External Warnings because the average speed for crossings with external warnings present was 55 kph and for crossings without external warnings was 40 kph. At 55 kph, 159 m is driven in 10.4 s, and at 40 kph, 118 m is driven in 10.6 s. A repeated measures t-test on participant time responses for External Warnings was not significant ($t(19) = 2.05$, $p = .06$). Therefore, as participant time responses did not differ for External Warnings, the third hypothesis was not supported.

For the four crossings with external warnings, the R-AWS were at 11.1 s and X-pavement markings at 9.6 s. The mean time response was 10.3 s (SD = 0.96), which is closer to the crossings than existing post-mounted R-AWS (M = 11.1). Using time responses, IRW would activate after R-AWS are passed by drivers, and thus, IRW would not act as pre-cues to R-AWS. Therefore, the fourth hypothesis was also not supported.

Figure 4. IRW onset distance for season conditions



The mean response for the 16 railway approaches was 10.3 s. The responses ranged from 8.8 to 11.7 s, and were highly consistent, as indicated by low standard deviation (SD = 0.9625), through the 16 approaches. These results are similar to those found in the pilot study (M = 9.1 s, SD = 1.31) (Chugh & Hove, 1997). Table 3 shows that approach speeds for the 16 railway crossings ranged from 30 to 65 kph. The decision sight distance criteria state that, for this range, the recommended time to respond is 10.2 to 14.0 s (Olson, 1996).

A stem-and-leaf/histogram test for extreme values within each of the 16 video clips showed that of 320 total responses, 10 were extreme. The 10 responses represented 3.125% of the data. Seven were 3 SD's above the mean and 3 were 2 SD's below. There was no systematic trend in the extreme points within the 16 video clips. There was one participant who consistently responded closer to crossings than other participants. The experimenter noticed this, and explained the protocol twice. The task was understood properly. As no clear reason was observed for the extreme points, they were discarded.

Exploratory Analysis

Differences between curved and straight approach responses were analyzed. Recall that of the 16 driver approaches to railway crossings (both summer and winter), 12 were straight and 4 were curved. This imbalance is because financial and time constraints limited the videotaping of more curved approaches. A repeated measures t-test showed a significant difference between the two approach types ($t(19) = 4.712$, $p < .001$). Curved approach responses were about 1 s longer than straight approaches (M = 11.2 s, SD = 1.97 versus M = 10.0 s, SD = 1.50 respectively). This result confirms pilot data (i.e., straight M = 8.7 s, curved M = 9.1 s) that roadway geometry is an important contributor to when advanced hazard information should be presented. Thus, providing advance information about an intersection that is not in drivers' immediate field of view is important.

Asymmetric Transfer Effects

As each participant viewed clips of railway approaches in both summer and winter conditions, there was the possibility that asymmetric transfer/carry-over effects occurred from one viewing to the next (e.g., Poulton, 1982). The order in which the 16 clips were edited together was randomized, and each participant was assigned to one of five orders. Participant responses to first views of railway crossings, be they summer or winter, were compared to the second views. A difference score was computed for each individual. A score of zero would indicate that the order in which crossings were viewed did not effect the responses. Nine of the 20 participants had scores below zero (range from -0.13 to -1.63), one had score zero, and the remaining 10 were above zero (range 0.25 to 1.13). The sum of the 20 difference scores was exactly zero. As the sum of difference scores was zero, it is concluded that asymmetric transfer/carry-over effects did not occur.

Study 1 : Discussion

None of the four hypothesis were supported. Participant responses were not influenced by gender, season, or by the presence of external warnings. The subjective responses suggest that IRW should come on once existing post mounted R-AWS are passed. Exploratory analysis show that curved approaches are deemed to require more time than straight approaches. In other words, when visibility to crossings is limited, more time is necessary to perform actions to avoid potential hazards. The variable of approach geometry is one of several variables that can effect visibility. Vegetation and buildings close to intersections are other sources of reduced visibility. Therefore, earlier information provision about upcoming hazards at curved approaches is recommended in designing IRW systems.

A consistent result has been the time prior to crossings the IRW should appear. Pilot and Study 1 results place the information in the range of 9.1 s to 10.3 s. As only younger individuals (approximately 20 years old) participated in Study 1, this range is suitable for them, and may be short for other age groups (especially for those above 65 years). The convergence of the 10-second mark with descriptions of decision sight distance (see Olson, 1996) is clear. As vehicle speed increases, then the distances, in advance of train crossings, at which warnings activate should also increase. For example, if warnings appear at 10 s before crossings, then at 40 kph the IRW would appear at 111 m, but at 60 kph the IRW would appear at 167 m. Warnings at 10 s give younger motorists enough time to perform decision sight distance tasks specified by Olson (1996) and Alexander (1991).

The Minnesota Department of Transport has recently developed an in-vehicle signing system for school buses at railway-highway grade crossings. The onset of the warning signal that informs bus drivers about approaching trains is at approximately 92 m (300 feet) before the crossing (Osemenam, personal communication, June 16, 1998). A motorist traveling at 45 kph would require 7.4 s to reach the crossing (from 92 m) and at 35 kph would require 9.5 s. But if a motorist is traveling at 65 kph, as in the summer approaches at the 52 Street at 50 Avenue SE (Northbound) crossing (see Table 3), then warning notification at 92 m would be equivalent to 5 s. If warning time of 10 s is desired, as shown by Study 1 results, then warning notification should be further away, at approximately 181 m. Alexander and Lunenfeld (1984) place decision sight distance of a vehicle traveling at 65 kph at 229 m or equivalent to 12.7 s of warning.

Motorist Expectancy

Expectancy is a driver's readiness to respond to situations, events or information in predictable ways (Alexander and Lunenfeld, 1975). A motorist's expectancy can be either brought to the driving task from past experiences or can be structured on the basis of the current operating environment. They play a large role in motorists' behaviour at railway crossings (Lerner et al., 1990). These include the type of TCDs (e.g., flashing lights and ringing bells) that are used at crossings, advance warning time as trains approach crossings, length of delays as trains travel through crossings, probability of being able to cross the tracks safely before trains arrive, and likelihood that trains will be coming toward crossings at different times. A driver's expectancy at a crossing site influences what is seen, how it is interpreted, the risks the driver is willing to assume, and what response alternatives are thought to be in appropriate (Lerner et al., 1990).

If onset of IRW for approaching trains is at 10 s, then the expectancy that crossings are always 10 s away eliminates the need to detect if there are crossings (because IRW not only tell drivers that trains are coming but also that crossings are up ahead), and reduces the need to detect exact locations of the crossings (as after onset on IRW crossings will always be 10 s away). This would be if drivers drove at speeds which were used to calculate distances at where the IRW activate. IRW also eliminate the need to detect if crossings are actively controlled (as IRW in themselves mean that crossings are actively controlled). Therefore, use of IRW would not only identify approaching trains, but would also simplify detection of other cues at railway crossings.

Implications for Design

The data collected in Study 1 were subjective responses by a small sample of participants watching video recordings of approaches to a limited number of railway crossings. Therefore, the results and recommendations should be interpreted with caution and need to be verified by field experiments. Drivers seem to want in-vehicle train warning information at the same time (10 s) regardless of whether other externally placed warnings are present. While the study did not explicitly ask participants if certain locations should have IRW, all volunteered to say that IRW were a "good" idea. The desire to have all crossings so equipped needs further empirical and feasibility testing.

Limitations of Present Study

A number of methodological concerns must be addressed. First, the time resolution with which data were recorded for Study 1 was 0.5 s. This is because the VCR's timer displayed time in seconds and not milliseconds. Because video is recorded at 30 frames per second (fps), participants may have chosen locations closer to the next second than the preceding second. This was not apparent on the timer. Therefore, any response from the 16th to the 29th frame (i.e., 0.5 s) would be indicated and recorded as the preceding second instead of the following second. A related concern is that vehicle speed during winter taping was slower than in summer taping. This limitation could not be avoided, given the environmental and financial constraints. Due to similar constraints, only eight crossings (six straight and two curved) were videotaped.

Finally, the study did not address when in advance of trains' approaches, signals should be transmitted to oncoming vehicles. The study of the two approaches (both of vehicles and trains) taken together is essential for the development of IRW. How IRW will effect the interaction between trains and vehicles will assist in understanding how drivers currently approach railway crossings and, at times, decide to disregard warnings in front of approaching trains (see Leibowitz, 1985).

External Validity

The findings from Study 1 are based on subjective participant choices using videotaped approaches. Participants were alert to the nature of the study and thus, were aware that they were selecting warning onset times for rail crossings. Therefore, the degree to which strong conclusions can be made about the safety benefit of IRW is limited, especially for drivers who are not aware that they are approaching railway crossings. While the data do converge with decision sight distance calculations, objective measures and larger samples (of participants and crossings) are required before the issue of placement of IRW is resolved. This study can, however, serve as a pilot test for field experiments that may place participants in real situations with real approaching trains.

Areas for Future Research

The 10 s value, selected by motorists of 18 to 22 years, may be somewhat short for older drivers. Future research could address under what circumstances more time is required (e.g., inclement weather). The existing advance railway crossing warning, R-AWS (code WC-4), which is a pictorial representation of a roadway intersecting a railway (CUTCD, 1976), can be used as an IRW appearing in a HUD. While the match between what the R-AWS actually is and what it is intended to tell the driver is questionable, the fact that it is recognized as part of a set of signs already installed at crossings makes it a good choice for continued use. Another important research question is whether warnings inside vehicles are more easily detected than warnings outside vehicles, and if so under what conditions. Finally, a critical issue with IRW is the effect of unreliable warnings on motorist performance and subjective trust in the warnings. This issue provided the motivation for Study 2.

Study 2

Traffic Control Devices (TCDs) not only inform drivers about hazards, they also form future expectancies. When there is a match between presence of a warning sign and the presence of a hazard, then the value or credibility of the sign increases. If, however, drivers see warning signs and select appropriate speed and paths to avoid upcoming hazards, and the hazards are not there, then the credibility of warning signs is reduced. This reduction in credibility effects future interaction of drivers with similar signs (Lerner et al., 1990). An example is a construction sign where there is no construction going on. In this case, the false alarm may lower driver credibility of, and trust in construction signs. While drivers may respond appropriately during the events for which the false warnings appear, responses to subsequent warnings, which may be valid, will reflect the reduction in credibility. When expectancies are violated (for example, by false alarms), then decision-making time and the likelihood of making errors increase (Olson, 1996).

When excessive false alarms occur, then users start to *disuse* (i.e., not use) systems (Parasuraman & Riley, 1997). While these statements about the effects of reduced reliability have been substantiated by research in other domains, very little is known about the effects of failures (both false alarms and missed signals) on performance with IRW. The effects of nuisance alarms, false alarms, and reduced reliability of any in-vehicle warning system will be crucial for future deployment, and more research in this area is required (Hancock, Parasuraman, & Byrne, 1995, p. 347). A recent review of human factors at railway crossings gave a *high* need for study of the effect of false alarms on driver behaviour (Carroll & Helser, 1996). The possibility of IRW failures does exist at railway crossings (Tardiff, Parviainen & Ede, 1996).

Some research has been done with reduced reliability in other domains that use new technology to automate operations (Lee & Moray, 1994; Muir & Moray, 1996), to provide navigation (Kantowitz, Hanowski & Kantowitz, 1997), collision warning (Lerner, Dekker, Steinberg & Huey, 1996) and radar patrol information (Nohre, MacKinnon & Sadalla, 1998). These studies have measured subjective trust or annoyance with systems when they fail. The studies are reviewed, and knowledge about failures (both false alarms and missed signals) that is generic is compared.

Trust Reduction with Automation Failures

Process control functions are becoming increasingly complex, and can be performed either manually by operators or automatically by other systems (Muir, 1994). In the automation mode (e.g., a system performs the process control functions in a nuclear plant), the task of an operator is to monitor proficiency of the system and to take manual control if the automation malfunctions. An automation can be expected to be competent in performing routine tasks but not be able to interpret or respond to unfamiliar situations. When an automation fails, then the task of an operator is to judge the level at which it failed and then either initiate a rule- or a knowledge- based response depending on his/her familiarity with the situation (see Rasmussen, 1983).

Muir (1994) hypothesized that when an automation fails, then operators' trust in it would decline. She likened trust in social settings to trust placed in machines, and proposed a model based on Rempel, Holmes & Zanna's (1995) hierarchical, sequential three-stage model. The hierarchical model focused on the development of trust between individuals, and started with development of *predictability*, followed by *dependability*, and culminating in *faith*. Muir (1994) suggested that these stages are relevant in development of operator trust in human-machine systems. The stages also occur in a fixed order, as in social trust, for trust in machines.

Aspects which influence predictability include the degree to which actions of the referent (i.e., individual or machine) vary, the degree of transparency which allows the operator to see how the individual or machine works, and the stability of the environment in which the referent operates. Dependability is developed through repeated exposure to the referent, especially during high demand conditions in which the referent's ability to perform is tested. Finally, faith is defined as a leap beyond the highly dependable and predictable behaviours of the past.

Using a milk pasteurization process control plant, Muir and Moray (1996) tested the effect of automation performance on the development of trust through the three sequential stages in two experiments, both with 6 participants each. Each participant completed 36 questionnaires through 9 hours of testing. As the rate of automation failures increased, trust in the automation decreased. When the failures either stopped or the rate at which they occurred decreased, then levels of trust rose, but at a slower pace and did not reach the initial levels.

After the automation initially failed and the subsequent reduction in trust, Muir and Moray (1996) manipulated how the automation's reliability varied, either in a variable or a constant manner. When the errors occurred consistently (i.e., constant error), then trust in the automation rose after initial declines. When failures happened in a variable manner, then trust remained low. Muir and Moray (1996) concluded that, "operators seem to be willing under some circumstances to adjust their trust in light of further evidence" (p. 441), revealing the dynamic nature of trust. Changes in trust changed how operators interacted with the automation. The more operators trusted the automation, the less intensely they monitored it. Conversely, the less they trusted the automation, the more they monitored it.

Muir's research was followed up by Lee and Moray (1992; 1994) who included the variable of manual or automatic control (i.e., by making it possible for the system to be controlled in both modes). Lee and Moray (1994) showed that operators' self-confidence in their ability to perform system functions manually mitigated the use of the automation. When operators were confident, then reductions in trust (induced by occurrences of system failure) resulted in inclinations to intervene and manually operate the system. However, when self-confidence was low, then the automation was relied upon to perform the functions even during low trust situations. Lee and Moray (1992) also tested for the development of trust through stages of predictability, dependability and faith. They found that faith, which was interpreted as a deeply held belief about the capabilities of a system, varied from participant to participant. Of the other two constructs, predictability was found to be dependent on observable behaviour, and dependability on the occurrence of faults.

In summary, these experiments have shown that reduction in the reliability of an automation reduces trust in it. When trust is reduced, then an automation's performance is increasingly monitored and control is taken when confidence in one's own ability to perform the same tasks as the automation is high. When the failures are predictable and consistent, then operators' trust of the system seems to grow after an initial decline, more so than when the rate of failures is varied and unpredictable.

Driver Annoyance with False Alarms

An auditory false alarm field experiment was performed with a collision avoidance system simulation, and responses were measured using daily and weekly questionnaires, over a 9 week period (Lerner et al., 1996). Noticeability (i.e., the degree to which warnings were noticeable), annoyance (i.e., the degree of annoyance), and acceptability (i.e., the rate at which the warnings occurred was acceptable) ratings were obtained. Bonus incentives were used to motivate participants to respond to true alarms and not to false alarms. Four dollars were paid to drivers each time true alarms were identified and one dollar was taken away each time responses were made to false alarms. Participants confirmed the presence of a hazard by looking at an indicator light on the far right of the dashboard. When the light was on, the hazard was present, and when it was off, the hazard was not present. Three of the four possibilities in the signal detection matrix were included in the test (i.e., hazard and warning present, hazard and warning absent, and warning present with hazard not present). The possibility of a missed signal (i.e., hazard present, warning absent) was not included, which is unusual considering that this is the most critical failure type.

False alarm rates of four per hour (acoustic alarm) and one per hour (verbal alarm) were deemed unacceptable by a sample of 15 participants. Less frequent acoustic warning rates of one per hour, one per 4 hours and one per 8 hours were not found to be as bothersome as higher false alarm rates. Lerner et al. (1996) identified two problems that result from false alarms. They are, "annoying, resulting in poor product acceptance, reluctance to use the product, or intentional defeating of the system" (p. 1), and they, "degrade the user's response, resulting in slower and less reliable reactions to valid warnings" (Lerner et al., 1996, p. 1).

Trust in Navigation Systems

Kantowitz et al. (1997) used three reliability levels (100%, 71%, and 43%) to test for differences in trust and use of a navigation system. Each of 48 participants was given four trials, and each trial consisted of traveling from a starting point to a destination point using seven links. At the start of the trials, participants marked, on a paper map, routes they would normally take. Before any of the 28 links in the study were chosen, participants obtained traffic congestion information about the links from the navigation system. If a link was deemed adequate by a participant, then it was chosen, or else another was selected. The traffic congestion information was either accurate or inaccurate. Once a decision/selection on a link was made, a video showed a drive through that link. The video either confirmed the traffic congestion information or contradicted it. If inaccurate links were chosen which increased travel time, then participants were penalized between $0.05 to $2.14. At the end of each link and video, a trust questionnaire was administered. The questionnaires asked participants to rate their trust in the system, and their self-confidence in the ability to find routes to the destinations without assistance of the navigation system.

The first two trials were 100% reliable and the other two were either 71% or 43% reliable; reliability was a between-participant variable. For the 71% reliable trial, the second and sixth links were always inaccurate, and for the 43% accurate trial the second, fourth, fifth, and sixth links were always inaccurate. Two of the four trials were for a familiar city and the other two were for an unfamiliar city (for participants from Seattle, Washington); familiarity was a within-participant variable. Kantowitz et al. (1997) used both types of incorrect matches that can take place with traffic congestion information. Inaccurate information was either harmful (i.e., traffic on link was reported to be light but was actually heavy) or harmless (i.e., traffic was reported to be heavy but was actually light). A number of specific scenarios were not addressed by Kantowitz et al. (1997); for example, what a user would do if he/she received erroneous information and followed it but ended up going the wrong way on a one-way street.

Participants placed highest trust in the navigation system when information was most reliable (i.e., 100%). The lowest level of trust was when the navigation system was least reliable (i.e., 43%). No statistically significant difference in trust was found between 100% & 71% reliable information. This led the researchers to conclude that navigation information that was 71% reliable would still be used by drivers. A significant difference in trust was not found between the two types of inaccurate information, each reducing trust to the same degree. Because trust was questioned after each link, participants referred to the last link in estimating their trust in the system. Therefore, when a link was accurate, trust was high and when it was inaccurate, trust was low.

Objective measures of penalty cost, convergence to marked map routes, and number of times the system was queried were analyzed. Penalty costs were higher when information was less accurate. This is simply a characteristic of the experiment; when the information is less accurate, then more inaccurate links are going to be chosen and more penalties are going to be imposed. When the trial was in a familiar city, then penalty costs were higher than when the trial was in an unfamiliar city. This suggests that participants chose more inaccurate links in the familiar city than in the unfamiliar city. Convergence to marked routes was higher when information was accurate. More query requests were made when navigation information was inaccurate.

In summary, as inaccurate information was presented, reliability of the system decreased, resulting in decreased trust. As trust decreased, number of queries of the system increased, supplemented with a divergence from pre-determined routes.

Effects on Driving Speed of False Radar Warnings

Nohre et al. (1998) asked participants to indicate hypothetical speeds on route to an airport to catch a departing flight. In written scenarios, participants were given 24 road descriptions, which contained either true warnings about placement of photo radar on the roadways, false warnings, or no warnings. Participants who received true warnings, indicated the slowest speeds of all three groups. Those who received false warnings initially indicated speeds similar to those of true warning participants, but as the number of false alarms increased, speeds also increased. Towards the latter part of the 24 road descriptions, those who had received false warnings were driving as fast as those who received no warnings at all. Nohre et al. (1998) concluded that, "results indicate that repeated false alarms decrease self-protective behaviour to a point suggesting that the warnings are eventually ignored" (p.1627).

Summary

While this review of effects of false alarms, automation failures, and inaccurate information on trust and performance is limited, it does reveal certain trends. When failures occur, they violate operators' expectancies of how systems should operate. The results of these violations are reduced trust and a desire in operators to find other ways of performing system functions. Automation failures make operators want to shift the operation from automatic to manual control (Lee & Moray, 1994). Operators are inclined to monitor an automation more when it fails (Muir & Moray, 1996). Driver queries of a faulty navigation system increase after reliability declines (Kantowitz et al., 1997). Drivers get annoyed when failures occur too frequently and within short time frames (Lerner et al., 1996). Hypothetical vehicle speed increases once warnings are known to be false (Nohre et al., 1998).

The salience of these findings across domains suggests that effects of failures may generalize to in-vehicle warning systems. These effects can be measured using subjective (e.g. trust, annoyance) and objective (e.g., monitoring, number of queries, hypothetical speed) dependent variables. Study 2 is a replication of previous research, testing the effects of reduced IRW reliability. IRW failures were designed to violate established expectancies, and subjective measures of trust and dependability, and objective measures of perception response time (PRT) were used. PRT was defined as the time it took a participant to respond by pressing a brake after onset of an IRW.

Present Study

The purpose of Study 2 was to determine the effects of reliability and different types of failures on performance (i.e., PRT) and trust in an IRW system. Two reliability levels were chosen, 83% (the appearance of an IRW correctly matched the presence of a train 83% of the time) and 50% (the appearance of an IRW correctly matched the presence of a train 50% of the time). Reliability levels of 83% and 50% were chosen because the experimental design (described below) allowed for these levels. There are three types of failures that can occur with IRW (see Table 5) (Tardiff, Parviainen & Ede, 1996). The first type, termed False Alarm 1, occurs when an IRW appears and there is a crossing up ahead, but no train arrives. The second type, False Alarm 2, occurs when an IRW appears and there is no crossing up ahead. The third type of failure, Missed Signal, takes place when an IRW fails to warn a train's arrival. Reliability and Failure type were presented as between-participant variables in Study 2.

Participants drove, using an accelerator and a brake simulation, through a number of approaches to railway crossings. In some approaches, auditory and visual IRW warned participants about approaching trains. Appropriate behaviour was to press the brake to slow or stop the car in the low-fidelity simulator. A video clip of a train at a railway crossing was added at the end of the approaches.

Table 5

Types of IRW failures

Failure Type	**Warning (Present / Absent)**	**Crossing (Present / Absent)**	**Train (Present / Absent)**
False Alarm 1	Present	Present	Absent
False Alarm 2	Present	Absent	Absent
Missed Signal	Absent	Present	Present

After gaining practice (at 100% reliability) with the meaning of the IRW, the appearance of a train, and responding by slowing down or stopping, participants experienced one of two reliability (83% or 50%) conditions. IRW failures differed by type (i.e., False Alarm 1, False Alarm 2 or Missed Signal). Following the failures, reliability returned to 100%. PRT and trust were measured before, during and after failures. Participants were rewarded for appropriate responses. Inappropriate responses were penalized.

Hypotheses

1. Decreases in reliability would effect PRT. Drivers would initiate braking later to warnings after experiencing the failures.
2. System trust would decrease after failures.
3. Trust would decline more in the 50% reliability condition than in the 83% reliability condition.
4. Trust would return to levels afterwards which were lower than baseline trust levels.

Study 2: Method

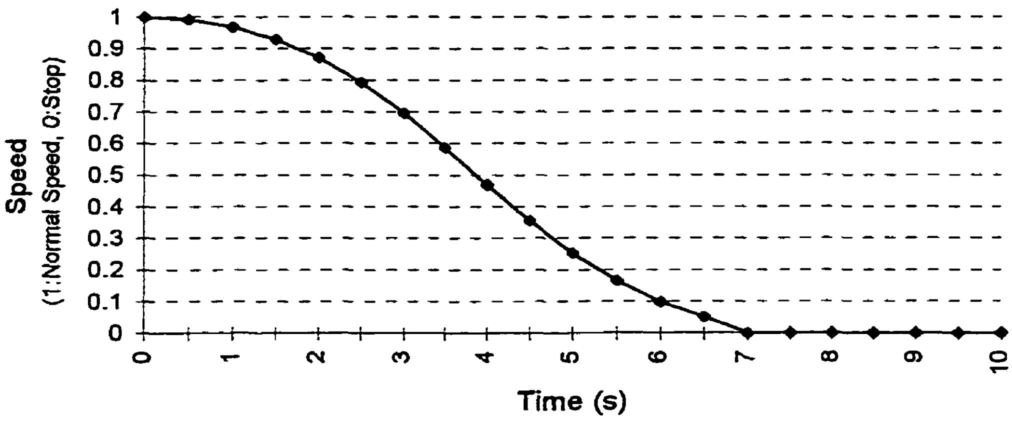
Participants

Thirty-six volunteers (18 male and 18 female), aged 18 to 26 years (M = 22), all with valid drivers licenses participated in Study 2. They had, on average, 5 years of driving experience and drove about 15, 000 km/yr. Twenty-eight of the 36 drivers had corrected vision lenses (i.e., glasses or contacts) which were worn during Study 2 testing, with a mean corrected visual acuity of 20/18 and normal contrast sensitivity, tested at 36 cm. All received a minimum of $5.00 for participating and up to $10.50, depending on the appropriateness of their responses.

Materials

Low-fidelity Simulator. Participants were seated in front of a simulator that had a brake, an accelerator and a steering wheel at positions similar to those found in a 1987 Chevrolet Sprint (see Appendix C). In front of them was a projection screen showing video clips. Pressing the brake caused the speed of the video clips to slow down and stop, as per a non-linear deceleration curve (see Figure 5) (Kaiser & Phatak, 1993).
Four deceleration curves, which differed in the rate in which the video decelerated and stopped, were presented to 6 pilot participants. The pilot participants chose the curve depicted in Figure 5 as the one thought to approximate braking. When the brake is

Figure 5. Brake deceleration in 7 seconds



continually pressed, it takes 7 s for the video to stop. The braking pattern, for the curve in Figure 5, is initially soft, becomes hard after 2 s and then eases off at approximately 4 s. Different deceleration styles (e.g., hard, soft braking) are used as strategic means to respond to situational constraints (i.e., vehicle speed, available stopping distance, and road surface wetness).

Pressing the accelerator gave the participants the ability to select three levels of speed at which the video could be played. An increase in the rate at which the video was shown approximated vehicle velocity increases. The levels were: *Normal* (i.e., the same rate/vehicle speed at which the video was recorded), *Medium* (i.e., 1.25 times the recorded speed), and *Fast* (i.e., 1.5 times the recorded speed). Mechanically, as the accelerator was pressed a cardboard attachment inside the simulator rose applying a current to the game port of a Gateway 2000 Pentium™ computer, on which the videos were stored. The video was then played at normal speed. An infra-red light was passed through the openings of the cardboard attachment to register increases and decreases in the accelerator pedal position and apply corresponding voltages to the game port. Additional depressions caused video playback speed to increase to medium and fast speeds. If the foot was taken off the accelerator altogether, then video speed declined less sharply than the deceleration curve in Figure 5. This reduction in speed was designed to replicate the effects of roadway friction. A schematic diagram of the brake and the accelerator is in Appendix C.

A steering wheel, a speedometer, an odometer and a fuel gauge were mounted on the simulator above the brake and accelerator. Participants were able to move the steering wheel around, and it was recommended that they do so when the roadways in the video curved, but turning the steering wheel did not cause a corresponding change in the view of the video. The steering wheel could be moved closer to the participant (somewhat like tilt steering) so that participants of different heights could comfortably place their hands on the steering wheel and touch the brake and accelerator with their right foot. The speedometer, odometer and fuel gauge were visible to participants, but were not functional.

In-Vehicle Railway Warnings. The IRW consisted of a visual iconic display on the windshield, similar to existing, externally post-mounted R-AWS, and an auditory bell emanating from the dashboard, similar to warning bells heard at railway crossings. The visual warning was chosen to be iconic rather than textual because research has shown that icons are more quickly identified than word signs (Dewar et al., 1976). The warning was inserted into the video approaches at a 2-dimensional location 4^0 below horizon and in the middle of the screen using Adobe Photoshop™ and Adobe Premiere™ (see Appendix C). This simulates the approximate location at which a HUD appears on the windshield. The image size was 45 pixels x 45 pixels, which was equivalent to 4 cm x 4 cm on a computer screen. The visual angle of the warning image, on the retina of a driver seated at 70 cm from the computer screen, would be 3.27^0.

An acoustic warning was chosen instead of a verbal warning because research has shown that acoustic warnings are better understood (Tan & Lerner, 1995) and are rated as less annoying by listeners than verbal warnings (Lerner et al., 1996). Verbal warnings have the tendency to confuse listeners. The clanging bells provided redundant information about approaches of trains. The auditory warning emanated from an Altec Lansing Multimedia (volume setting 7 treble -3) speaker located behind the steering wheel of the simulated vehicle dashboard. This represents the location to which driver attention is to be directed (Huey, Harpster & Lerner, 1997). The purpose of IRW is to inform drivers that trains are coming, and to direct attention to front areas of vehicles so that preventative measures by drivers can be undertaken. When detectors can accurately predict directions from which trains are coming, then acoustic warnings may be designed to emanate from those areas within dashboards.

As recommended by findings of Study 1, both visual and auditory warnings were presented simultaneously at 10 s in advance the crossings. The duration of the warnings was 2.5 s. Auditory warning of 2.5 s has been found to be long enough for detection and comprehension by drivers (Tan and Lerner, 1995). As the IRW were embedded in the driving images, participants driving at the fast rate (1.5 time rate of normal driving) would see and hear the warnings for 1.7 s. Similarly, at the medium rate the warnings would be present for 2 s. The change in rate of video flow not only changed the duration of auditory warnings, it also changed the pitch (i.e., frequency). Therefore, while the visual warnings looked the same at all speeds, the auditory warnings were heard at higher or lower frequencies dependent on the speed of video flow. The rise and drop in pitch did not, however, lessen the impact of the warnings on participants' understanding and responses to IRW. The change in frequency may, however, have been taken as a rise in urgency (see Haas & Casali, 1995), which would be of benefit as it would raise drivers' alertness towards train hazards.

Digitized Video. The analog videos used in Study 1 were digitized using Adobe Premiere™ at 30 fps at 320 x 240 pixels, using a Power Macintosh 8100/80 computer at the New Media Center in Department of Communications Media, University of Calgary. The approaches were composited and linked together into movies and IRW were inserted using Adobe Premiere™. Combining of the visual IRW with approaches was done using an Alpha Channel transparency matte in Adobe Premiere™. During the rendering process, the digital movies (saved as QuickTime™ files) were flattened (i.e., made cross-platform so that they could be played on a PC), keyed every frame (to allow manipulation at rates as high as 1/30 s), and compressed using an Intel Indeo™ Video R3.2 compression. This compression was necessary to keep the digitized video movies of manageable file size.

The digitized movies (38 in all) ranged in size from 75MB to 130MB, and were initially too big to be played on the Gateway 2000 Pentium™ because of low RAM

(32MB). The playing seemed to skip frames and disrupted the continuity of the video. Three measures were taken to reduce the disruption of continuity. The RAM was expanded to 96MB. The 2 hard drives of the Pentium were defragmented to allow clean throughput of files onto RAM. The movies were compressed again using Intel Indeo™ Video R3.2 compression, and the data rate of the movies was limited to 600 K/sec. These measures ensured that the movies played at or close to 30 fps.

Software. An Asymetrix ToolBook™ computer program played the movies according to information from the simulator controls (i.e., brake, accelerator). The output of the program (in ASCII format) contained participant responses, including brake press times (measured with the computer's internal clock), brake release times, and accelerator press and release times (for all three levels). A separate parsing program (also created using ToolBook™) allowed extrapolation of certain key fields (e.g., brake press times) for analysis.

A concern when using computers for experiments is time lag. A number of unavoidable events occur between activation of simulator controls (brake or accelerator) and their effects. First, signals travel along wires from the controls to the game port. Then, the game port reviews the signals, determines their source and transmits them to the ToolBook™ program. The ToolBook™ program then plays the videos at the rate indicated by the controls, and records the times at which control presses were made. Lags up to 5 milliseconds in the recording of responses and up to 15 milliseconds in change in speed were present in the study.

Movie Projection. The digitized movies were projected from the Gateway 2000 Pentium onto a projection screen (Da-Lite Picture King) using a Sharp Liquid Crystal Projector (Model No: XG-E3000U High Resolution). The screen was located 400 cm from the projector. Participants was seated 50 cm in front and 40 cm to the right of the projector, allowing the movies to be projected over/beside their left shoulders. Thus, participants were 350 cm from the screen. The movie images on the screen were 60 cm wide and 45 cm high. The movies were digitized at 320 pixels x 240 pixels, and the placement of the screen 400 cm away from the projection made the images 5.3 times the size of the computer monitor image. The visual warning was 45 pixels x 45 pixels, and it appeared as 21 cm x 21 cm on the screen. The visual angle of the warning image, on the retina of a driver seated at 350 cm from the projection screen, would be 3.44^0. The lower part of the image on the screen was 100 cm from floor level. This was the same distance the projector was from the floor. This allowed participants to maintain horizon line-of-sight at approximately the middle of the projection. The simulator was placed directly in front of the participants, and its position relative to the projector and the screen varied according to their comfort with pressing the brake and accelerator.

Rewards and Penalties. Participants were rewarded $0.50 for practicing safe behaviour which entailed slowing or stopping the vehicle (i.e., video) before trains arrived. If they slowed or stopped the vehicle and a train did not arrive, then they were charged $1.00 for causing congestion. If they failed to slow or stop, and a train came to the crossing, then they were penalized $2.00 for causing a crash. The reward/penalty values were chosen to maintain a consistent goal structure for participants throughout the study. The bonus incentives applied to only the third and fourth blocks (a block is explained below). Any money earned was above and beyond $5.00 each driver received for participating. Cash incentives, based on performance, have been used previously to maintain participant motivation in similar experiments (Lee & Moray, 1994; Lerner et al., 1996; Muir & Moray, 1996; Kantowitz et al., 1997).

Questionnaires. A two-page questionnaire (Appendix D) was administered at the beginning of the study to ascertain demographic information, visual characteristics as well as driving experience. In addition, this questionnaire described and tapped individual perceptions of trust and self-confidence. A single page questionnaire (Appendix E) was completed three times (after the second, third, and fourth blocks) to determine participants' trust in and ratings of dependability of the IRW system. Participants were also asked to estimate the relative number of times the trains appeared within a previous block. A three-page questionnaire (Appendix F) was completed at the end of the study that asked questions about design of IRWs, how the drivers would use the system given various degrees of reliability, and whether they would engage in a number of risky behaviours if late for an exam.

Study Design

Each participant interacted with the simulated system for about an hour. The terms used describing the experimental design structure are elucidated next.

Event. An event was an approach to a railway crossing or a driving scene with no crossing. Each event ranged in time from 20 to 40 s. At the end of an event, another was digitally attached.

Critical Event. A critical event was defined as an approach to a railway crossing during which a train was coming and/or an IRW was present.

Valid Critical Event. Critical events in which both an IRW and a train appeared defined a valid critical event. Ten seconds prior to the end of an approach an IRW was presented, and 4 s prior to the end of the approach a train clip pre-empted the rest of the approach. This train clip was 8 s long. The splicing together of approaches and train appearances showed a logical completion of valid critical events, and the montages were perceived as such (see, e.g., Hochberg & Brooks, 1978).

Invalid Critical Event. Critical events in which either an IRW or a train appeared defined an invalid critical event.

Block. There were 24 events in a block, and 6 (25%) of the 24 were critical events (the remaining 18 or 75% were benign events with no trains and no warnings). While the appearance of a train at a crossing is a rare event for any particular driver, a higher rate of occurrence (i.e., 25%) was required for this study because of financial and time constraints. In addition to railway approaches, other driving scenes were included. Four blocks were used and each block lasted about 12 min. Participants could take a break between blocks if they wished. Participants were required to drive through all 24 events in a block continuously, but could self pace the simulated drives by using the brake and accelerator. Self pacing allowed participants to control the rate at which the experiment progressed, which replicates driving in realistic traffic situations (Zaidel, 1991).

Procedure

Visual acuity and contrast sensitivity were tested using a Landolt C Near (36cm) Visual Acuity Chart and a Vistech Vision Contrast Test System (Type B) Chart at the start of the study. A two-page driving experience questionnaire was completed by each participant at this time. Then, the participants were seated in front of the simulator. Tasks to be completed and rewards/punishments were explained to them (see Appendix G for verbal protocol for Study 2). Several video approaches were shown to them to explain how to use the brake, accelerator and steering wheel. Examples of the IRW (with both visual and auditory modes) were also shown.

A participant's session began with two blocks of practice. Within each of these practice blocks all six critical events were valid. Thus, both practice blocks were 100% reliable (see Table 6), and responses to the practice blocks were taken as baseline. The appearance of the valid critical events varied randomly within the set of 24 events. As a number of critical events formed a block across time, the variable of Critical Events was taken as a within-participant variable, for the two practice blocks.

A 2 (Reliability) x 3 (Failures) x 6 (Critical Events) participant design was integrated into the third block. Participants were assigned to one of two conditions of Reliability (83% or 50%) and one of three conditions of Failures (False Alarm 1, False Alarm 2 or Missed Signal). Participants who were in the 83% reliable condition encountered five valid critical events and one invalid critical event, whereas participants who were in the 50% reliable condition experienced three valid critical events and three invalid critical events (see Appendix H). The third of the six critical events in the 83% reliability condition was the failure event. Likewise, for the 50% reliability condition the first, third and fourth of the six critical events were always the failure events. The reward/penalty structure applied only to the third and fourth blocks, and not the practice blocks. After completing the third block and the post-block questionnaire, participants were told their rewards or penalties for the third block.

A participant's session ended with a fourth block that had warnings which were 100% reliable. The order of presentation of the railway crossings was randomized across the four blocks (see Appendix I). At the conclusion of the four blocks, a post-experiment questionnaire was administered, and the participants were debriefed about the nature of the study and paid (see Appendix J for verbal protocols).

Table 6

Reliability levels for the four blocks for Study 2

Reliability Condition	**Block 1 (practice)**	**Block 2 (practice)**	**Block 3**	**Block 4**
83% reliability	100% 6/6 events valid	100% 6/6 events valid	83% 5/6 events valid	100% 6/6 events valid
50% reliability	100% 6/6 events valid	100% 6/6 events valid	50% 3/6 events valid	100% 6/6 events valid

Study 2: Results

Perception Response Time (PRT) Analysis

Participants were informed in the experimental protocol that when they saw an IRW, it was telling them that a train was approaching the crossing they were driving toward. In wanting to practice safe behaviour it would be prudent of them to slow the video/vehicle so that it could safely come to a stop before it reached the crossing. All IRW activated 10 s prior to crossings and at 4 s before the crossings, the train clips were inserted. Thus, the final 4 s of original clips were replaced with images of the trains. While brake presses throughout the blocks were recorded, only those within the 6 s from the onset of the IRW to the activation of the train clips were defined as PRTs and analyzed.

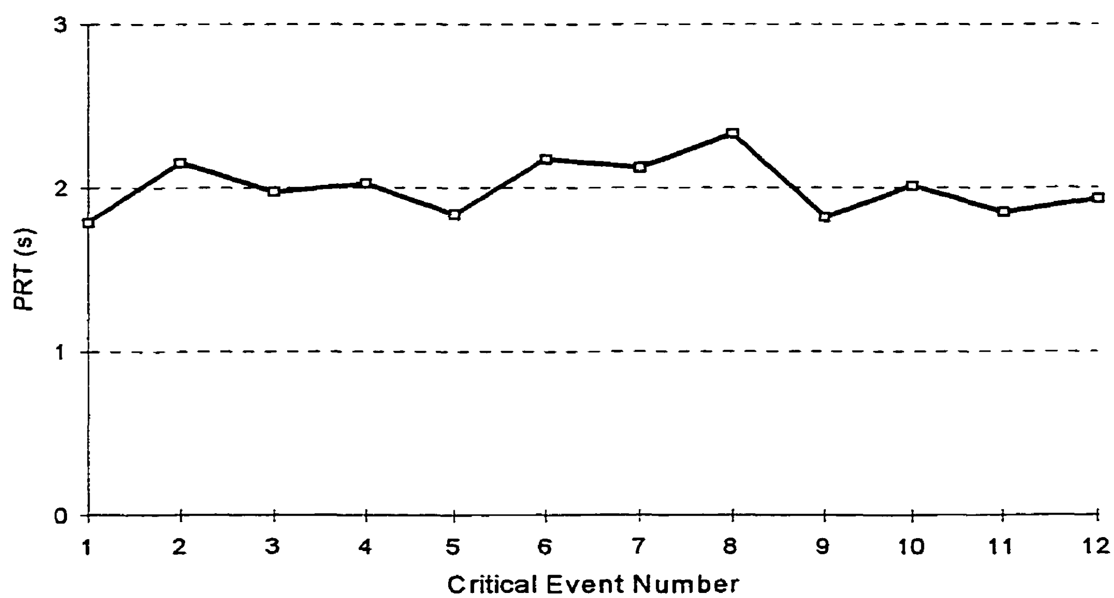
The reason that the brake presses were not analyzed during the last 4 s are two-fold. First, if a participant braked during the last 4 s then the inertia of the car would not allow it to stop safely before it crossed onto the tracks. Participants were told that if they hadn't pressed the brake (or the video hadn't dramatically slowed down) before the train clip activated, then that signified a crash. Secondly, as the train clip was separate from the original clip, it had to play at normal speed (regardless of the position of the brake or the accelerator). Any brake presses made before IRW came on were participants simply controlling flow of the videos and not necessarily responding to IRW.

As both practice blocks were 100% reliable, they were analyzed together using the single independent variable of Critical Events. There were 12 critical events across the practice blocks, and a repeated measures ANOVA was performed on PRTs of the 12 critical events. The assumption of sphericity was violated (Mauchly's W (65) = 0.024, p < .001). Because of the violation of sphericity, the Greenhouse-Geisser epsilon (Box's Ê = .646) was used to correct the degrees of freedom. The main effect of Critical Events was not significant (F(7.106, 248.71) = 1.945, p > .05). The lack of a main effect meant PRT responses did not vary significantly across time (see Figure 6).

The overall grand PRT mean, across the two blocks, was 1.80 s (SD = 0.80), which was used as baseline PRT. There was no difference (t(35) = -0.506, p = .616) between PRT to the first critical event in the first practice block (M = 1.82, SD = 1.1) and the last critical event in the second practice block (M = 1.97, SD = 1.07). By the end of the second block, participants had gained familiarity with the protocol and response expectancies.

While baseline PRT for Study 2 was at approximately the 85th percentile PRT in stopping sight distance situations (Fambro et al., 1998), it does adequately represent baseline PRT for this experimental setup. Researchers (e.g., Fambro et al., 1998; Olson & Sivak, 1996) suggest that brake press responses to unexpected objects is at about 1 s.

Figure 6. PRT for critical events through two practice blocks

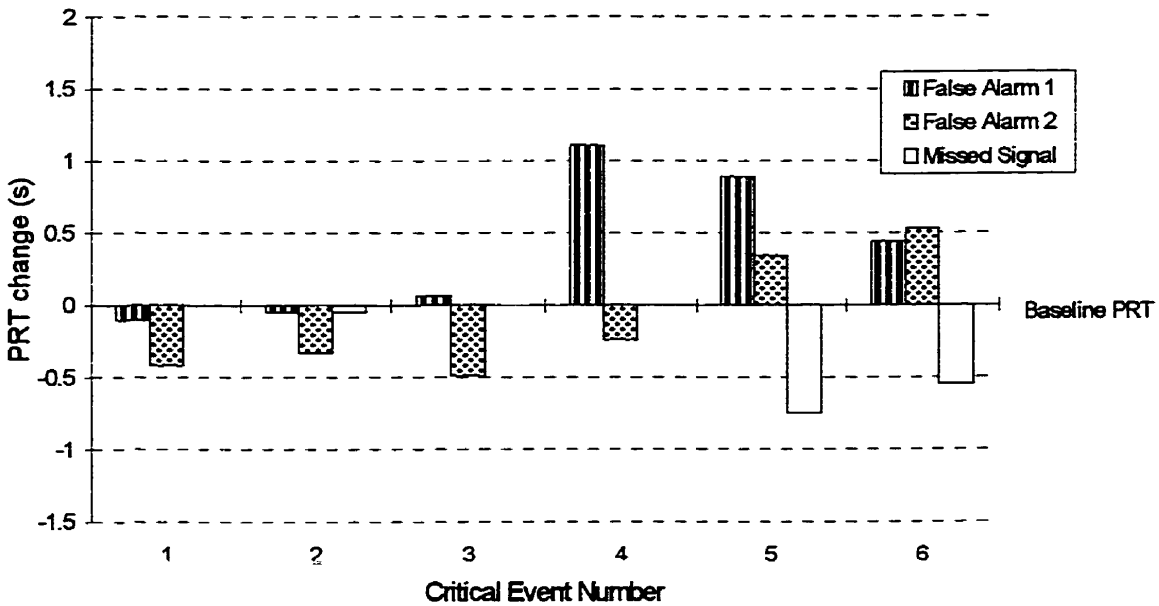


These responses are made in situations where objects encroach onto drivers' paths, and the task is to immediately bring vehicles to stop. In Study 2, however, participants were aware that the IRW would activate 10 s before the vehicles reached the crossings, and 6 s before the train clips would activate. Therefore, using self pacing (Zaidel, 1991) participants converged PRT at 1.8 s. Pressing the brake at 1.8 s would give participants sufficient time to comfortably stop before reaching the crossings, without creating sudden decreases in speed that are associated with hard braking at 1.1 s.

At the completion of the second block, participants were randomly assigned to one of two reliability and one of three failures conditions. While all participants in the 83% reliability, and False Alarm 1 and False Alarm 2 conditions could respond to all six critical events, the participants in the 50% reliability and Missed Signal condition could only respond to three of the six critical events; those being the second, fifth and sixth events. This is because when Missed Signals occurred, then the IRWs did not come on informing participants of approaching trains. No other cues were available to warn participants. Thus, analysis of PRTs for the third block was performed on valid critical events (i.e., on the fifth and sixth critical events) that followed the failures.

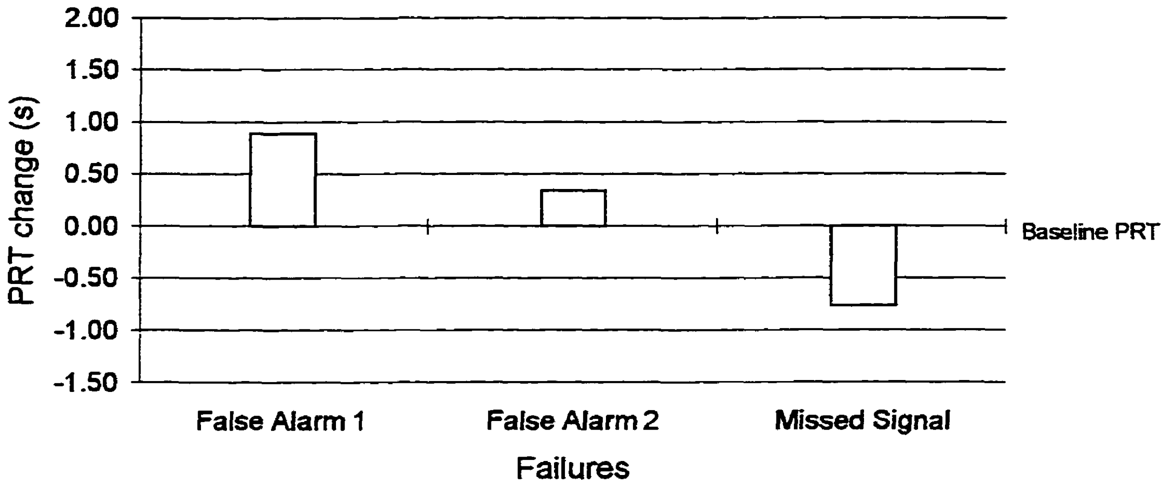
A difference PRT score for each participant was created by taking each event PRT and comparing it to baseline PRT (grand PRT mean from the practice blocks). The differences scores allowed for each participant to act as his/her own control. A mixed model ANOVA of Failures (3) x Reliability (2) x Critical Events (2) was done on the difference PRT scores for the third block. The three-way interaction was not significant ($F(2,29) = 0.485$, $p = .621$). All 3 two-way interactions were not significant; Critical Events x Reliability ($F(1,29) = 0.097$, $p = .758$), Reliability x Failures ($F(2,29) = 0.888$, $p = .422$), and Critical Events x Failures ($F(2,29) = 0.855$, $p = .436$). The main effects of Critical Events ($F(1,29) = 0.364$, $p = .551$) and Reliability ($F(1,29) = 4.048$, $p = .054$) were not significant. The marginal effect of reliability was two-tailed, and showed that changes in PRT for 50% reliability were greater than changes in PRT for 83% reliability. The main effect of Failures ($F(2,29) = 6.615$, $p = .004$) was significant. Figure 7 shows PRT change from baseline for all six critical events in the third block collapsed across reliability and divided by failure type. From the first to the fourth critical events, failures were taking place.

Figure 7. PRT change from baseline for third block collapsed across reliability



The main effect of Failures, for the fifth and sixth events, shows that PRT changed differentially depending on the type of failure that was experienced. Those who had false alarms increased their PRT (for False Alarm 1 by 0.89 s and 0.34 s for False Alarm 2) while those who had missed signals decreased their PRTs (by 0.77 s) (see Figure 8). PRT changes for False Alarm 1 ($t(11) = -2.281$, $p = .043$) and Missed Signal ($t(11) = 4.144$, $p = 0.002$) were significantly different from baseline ($M = 1.8$ s, $SD = 0.80$), but not for False Alarm 2 ($t(11) = -0.815$, $p = .432$). PRT changed to 2.54 s ($SD = 1.27$) for False Alarm 1, 2.37 s ($SD = 1.30$) for False Alarm 2, and to 1.26 s ($SD = 0.83$) for Missed Signal. Therefore, while PRT increased for those in false alarm conditions, so did response variances. PRT for missed signals decreased to levels described by Fambro et al. (1998), and Olson and Sivak (1996). As there was no significant effect of reliability, the responses were similar for those who had one failure and those who had three failures. However, PRT after False Alarms 1 and 2 in the 50% condition was higher than in the 83% condition, but PRT after Missed Signals was the same at both reliability conditions.

Figure 8. PRT change from baseline for third block for failures



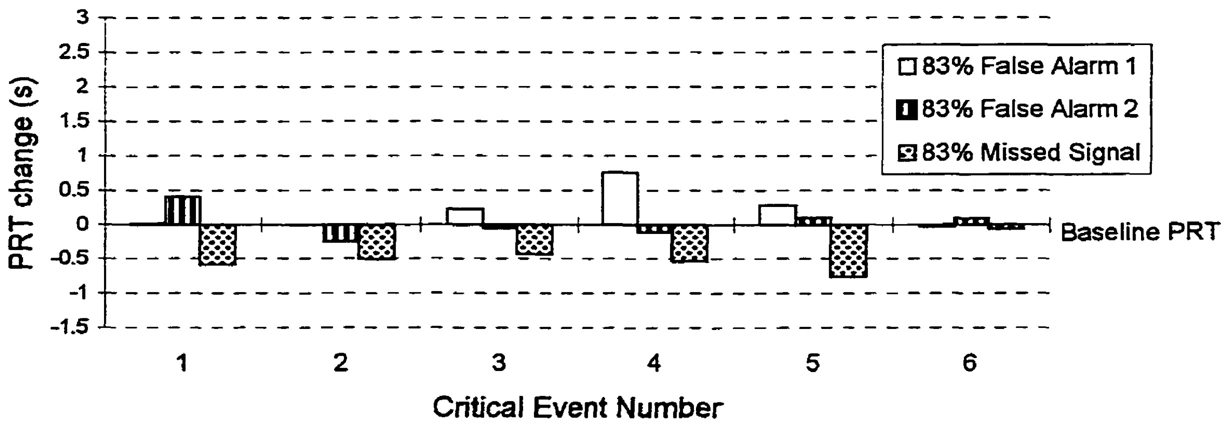
Two out of 108 critical events (1.85%) in the third block were missed by participants. Two participants intentionally did not press the brake after IRW came on. One of the missed events occurred in the fifth event by a participant in the 83% reliability condition who had experienced a False Alarm 2. The other occurred in the sixth event by a participant in the 50% reliability condition who had experienced three False Alarm 1. Both of these participants had been penalized previously for prior false alarm responses. In the real environment, these missed events would have resulted in crashes. These missed events were coded as 6 s, which is the highest possible PRT that could have been made by a participant. The reasons why missed responses were coded as 6 s was because removal of them would have reduced error degrees of freedom and certain cell sizes to below 5, and that they were part of driver behaviour and thus, should be part of the analysis. After each missed event, the 2 participants realized that crashes had taken place because they had failed to respond. One other participant's data, in the 50% reliability and Missed Signals condition, was discarded because PRT responses throughout the study were outliers. Time and financial constraints prevented replacement of this participant's data.

As all six critical events for the fourth block were valid, all six could be analyzed. A mixed model ANOVA of Failures (3), Reliability (2), and Critical Events (6) was done for the fourth block, again using PRT difference scores (computed using baseline PRT).

The three-way interaction ($F(10,145) = 2.098$, $p = .028$) was significant. Two two-way interactions were significant, Critical Events x Failures ($F(10,145) = 2.285$, $p = .016$) and Critical Events x Reliability ($F(5,145) = 9.563$, $p < .001$). The Reliability x Failures interaction ($F(2, 29) = 2.61$, $p = .091$) was not significant. In addition, two of the three main effects were significant; namely Critical Events (6) ($F(5,145) = 7.474$, $p < .001$) and Failures (3) ($F(2,29) = 10.487$, $p < .001$). The main effect of Reliability ($F(1,29) = 2.084$, $p = .16$) was not significant.

As the three-way interaction was significant it is interpreted first. Two two-way simple interaction ANOVAs were performed for 83% and 50% reliability conditions. The two-way simple interaction of Critical Events (6) x Failures (3) in the 83% reliability condition was not significant ($F(10,75) = 1.34$, $p = .226$), and neither were the main effects of Critical Events (6) ($F(5,75) = 0.71$, $p = .618$) or Failures (3) ($F(2,15) = 1.935$, $p = .179$). The lack of significance of the simple two-way interaction and the main effects shows that, for participants in the 83% reliability condition, PRT in the fourth block had returned to baseline levels. Figure 9 shows PRT responses across time and for the three failure types for the 83% reliability condition. As the omnibus three-way interaction was significant for the fourth block PRT difference scores, PRT for participants in the 83% reliability condition (collapsed across critical events and failures) did differ from those in the 50% reliability condition.

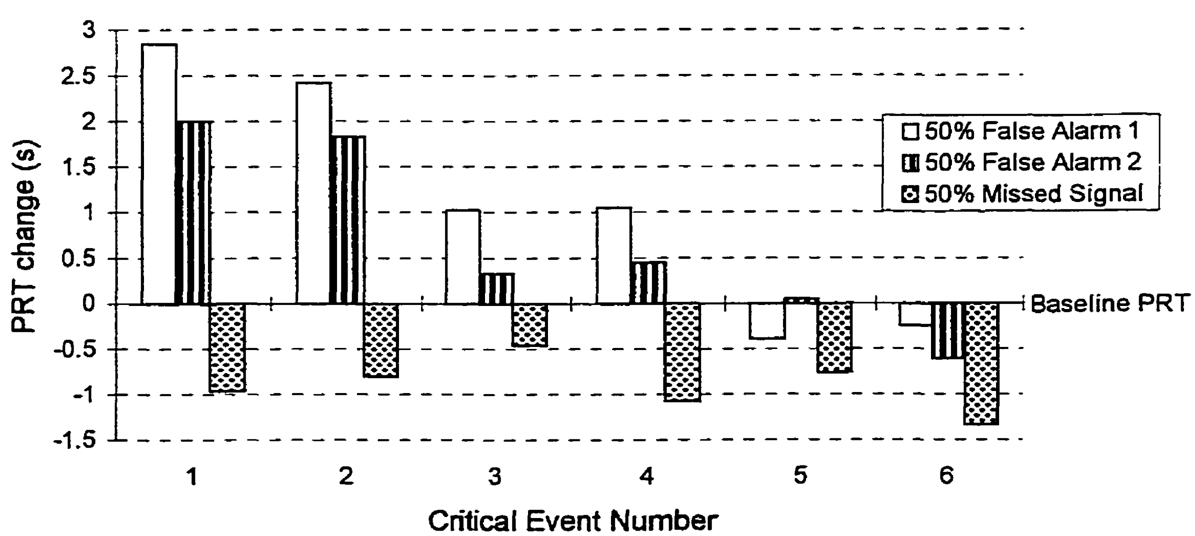
Figure 9. PRT changes from baseline for 83% reliability in the fourth block



The two-way simple interaction of Critical Events (6) x Failures (3) in the 50% reliability condition was significant ($F(10,70) = 2.301$, $p = .021$), as were the main effects of Critical Events (6) ($F(5,70) = 10.494$, $p < .001$) and Failures (3) ($F(2,14) = 8.589$, $p = .004$). The significance of the two-way simple interaction shows that for the 50% reliability condition PRT differences from baseline changed over time and by failure type in the fourth block (see Figure 10). Through the fourth block from the first to the sixth critical event, PRT for False Alarm 1 and False Alarm 2 decreased towards baseline, while PRT for Missed Signal remained consistently below baseline. Indeed, simple simple effects tests for Critical Events (6) for False Alarm 1 ($F(5,25) = 9.082$, $p < .001$) and False Alarm 2 ($F(5,25) = 3.943$, $p = .009$) were significant, but not for Missed Signal.

A differential lingering effect of the false alarms on PRT between participants in the 83% and 50% reliability conditions is evident. By the start of the fourth block, PRT for those in the 83% reliability condition had returned to baseline. In contrast, those in the 50% reliability condition took until approximately the fifth critical event in the fourth block to revert PRT to baseline. Therefore, it appears that the more false alarms drivers experience, the longer it takes them to return to pre-false alarm levels, when the

Figure 10. PRT change from baseline for 50% reliability in the fourth block



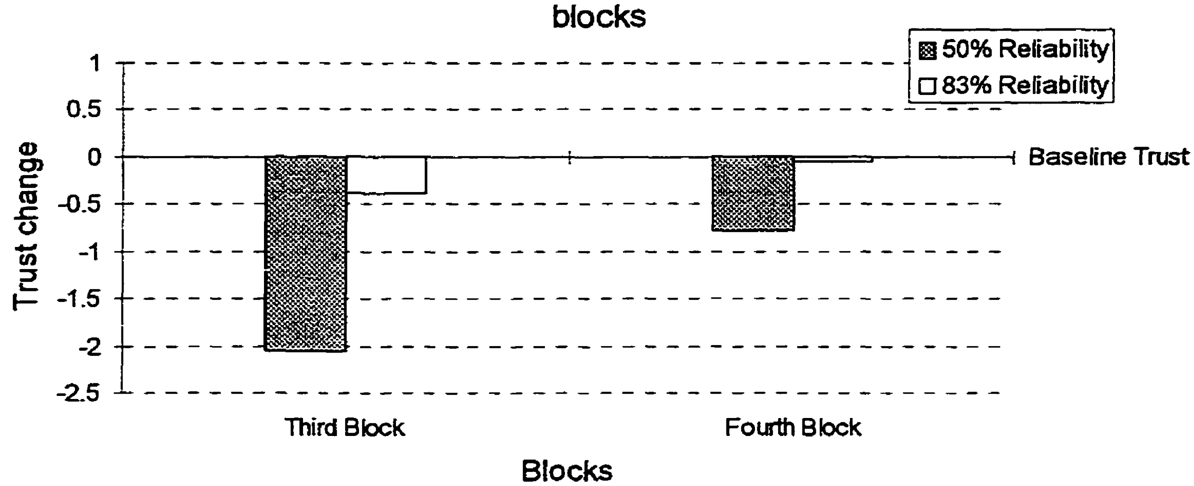
false alarms cease to happen. Those in the 50% reliability condition also received more penalties than rewards because of the three failures. PRT for those in the missed signals remained below baseline throughout, however, for both reliability conditions. PRT results have provided partial support for the first hypothesis. PRT was indeed influenced by failures, and increased for those in the false alarm conditions, but decreased for those who experienced missed signals.

Ten of the 216 responses (4.6%) made for the six critical events in the fourth block were missed events. Like instances in the third block, these missed events were because some participants intentionally did not press the brake after IRW had come on. Eight of the missed events were by participants in the 50% reliability condition and 2 were by those in the 83% reliability condition. Of these 10 missed events, 5 were for False Alarm 1 condition and 5 for False Alarm 2. No participant in the Missed Signal condition had a missed event. Recall that participants were penalized $1.00 for creating congestion if they stopped unnecessarily. All participants who had missed events had lost money in responses to previous false alarms. Like those in third block, these 10 missed events may have resulted in crashes in the real environment. Once the trains arrived, the participants realized that crashes had taken place. These responses were again coded as 6 s.

Questionnaire Results

For the trust dependent variable, difference scores were computed by comparing responses from the third and fourth blocks to the baseline measure (i.e., second (100%) block responses). Trust was measured on a scale from 1 to 7 with ratings of 1 for absolutely no trust and of 7 for absolute trust. For analysis of the third block, trust difference scores, a two-way ANOVA was performed, with Failures (3) x Reliability (2). The two-way interaction ($F(2,30) = 1.963$, $p = .158$) and the main effect of Failures ($F(2,30) = 0.319$, $p = .729$) were not significant. The main effect of Reliability was significant ($F(1,30) = 15.101$, $p = .001$) (see Figure 11, third block). Figure 11 and the main effect show that the trust differences from baseline differed between the two reliability conditions after the third block. Trust dropped by 2.06 units for 50% reliability from baseline and by 0.39 units for 83% reliability. This result supports the second hypothesis, that system trust would decrease after the failures. Trust for 50% reliability differed significantly from baseline ($t(17) = 7.2$, $p < .001$), but not for 83% reliability ($t(17) = 1.917$, $p = .248$). This result supports the third hypothesis that trust would decline more for 50% reliability than for 83% reliability.

Figure 11. Trust change from baseline for the third and fourth blocks



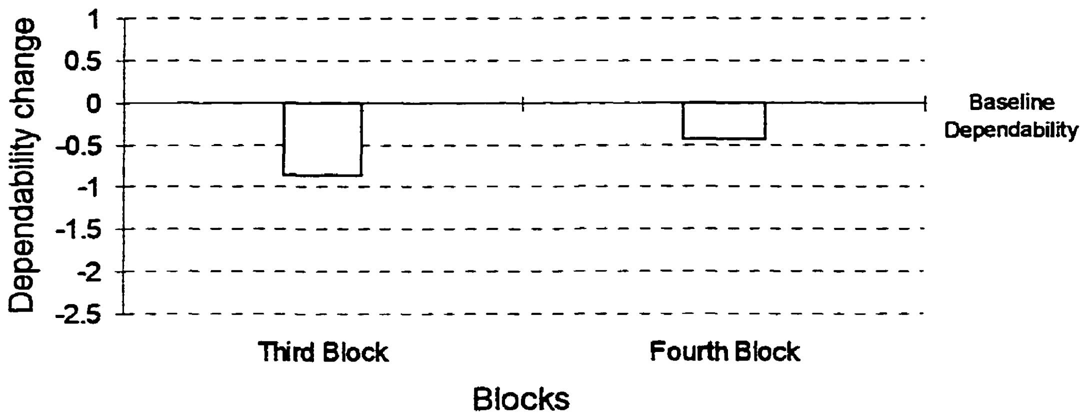
For analysis of the fourth block trust difference scores, a two-way ANOVA was performed, with Failures (3) x Reliability (2). The two-way interaction ($F(2,30) = 0.159$, $p = .854$) and the main effect of Failures ($F(2,30) = 0.124$, $p = .884$) were not significant. The main effect of Reliability ($F(1,30) = 3.975$, $p = .055$) was marginally significant (see Figure 11, fourth block). Trust had returned to baseline levels after the fourth block for the 83% reliability condition, but had not for 50% reliability. Trust remained 0.78 units below baseline for 50% reliability and 0.05 units for 83% reliability. Trust for 50% reliability differed significantly from baseline ($t(17) = 2.715$, $p = .015$). This result partially supports the fourth hypothesis that trust would return to levels lower than baseline after the return to 100% reliability.

These findings for trust changes in the third and fourth blocks are consistent with prior studies in process control (Lee & Moray, 1992; Muir & Moray, 1996), and ITS navigation systems (Kantowitz et al., 1997). The null effect of failure type suggests that the measure of trust was not sensitive to the type of failure (see Kantowitz et al., 1997 for similar findings). In other words, any violation of expectancy decreased trust. The use of failures as a between-participant variable, however, means that participants did not experience all types of failures, and thus were unable to compare between failure types.

A Pearson correlation test was performed using trust difference scores from the third block and PRT difference scores from the first three critical events of the fourth block. For False Alarm 1, trust after the third block was significantly correlated with PRT from the fourth block ($r = -.651$). The negative direction of the correlation means that as trust decreased PRT increased. A similar correlation result, however not significant, was found for False Alarm 2 ($r = -.434$). For Missed Signals the correlation between trust and PRT was low ($r = .073$), but was in the direction suggested by earlier PRT and trust analysis (i.e., as trust decreased PRT decreased). Therefore, changes in trust were related to changes in PRT, especially for false alarms. Baseline trust from the second block could not be compared to PRT from the third block because failures were taking place in the third block and, as mentioned earlier, some participants could not respond to all events in the third block because of the failures.

A measure of dependability of the IRW system was also obtained from participants at the end of second, third, and fourth blocks. Dependability responses at the end of the second block were taken as baseline scores (for 100% reliability). Dependability was measured on a scale from 1 to 7 with ratings of 1 for absolutely no dependability and of 7 for absolute dependability. Dependability difference scores were computed for the third and fourth blocks, using the baseline as comparison. For the third block, a two-way interaction of Reliability (2) x Failures (3) was not significant ($\underline{F}(2,30) = 1.627$, $\underline{p} = .213$). The main effects of Reliability (2) ($\underline{F}(1,30) = 1.806$, $\underline{p} = .189$) and Failures (3) ($\underline{F}(2,30) = 1.358$, $\underline{p} = .272$) were also not significant. The lack of significance means that any difference from baseline for the third block did not differ by type of reliability or failure. A paired t-test showed that the difference from baseline in the third block was significant ($\underline{t}(35) = 3.645$, $\underline{p} = .001$) (see Figure 12, third block). Figure 12 shows that dependability dropped significantly from baseline after the failures, by 0.86 units.

Figure 12. Dependability change from baseline for the third and fourth blocks



For the fourth block, a two-way interaction of Reliability (2) x Failures (3) was not significant ($F(2,30) = 0.09$, $p = .915$). The main effects of Reliability (2) ($F(1,30) = 0.796$, $p = .379$) and Failures (3) ($F(2,30) = 0.316$, $p = .732$) were also not significant. The lack of significance means that any difference from baseline for the third block did not differ by type of reliability or failure. A paired t-test showed that the difference from baseline in the fourth block was significant ($t(35) = 2.239$, $p = .032$) (see Figure 12, fourth block). Figure 12 shows that while dependability started to return to baseline levels, it remained significantly below baseline (by 0.43 units) after the system had returned to 100% reliability. As the measure of dependability did not vary with reliability (i.e., no interaction or main effect), it shows that participants decreased their ratings of dependability of the IRW system during times of failure. In contrast, while trust also decreased during failures, it decreased more so at lower reliability. Participants were also asked to estimate how often trains appeared during the blocks. In the second and fourth blocks, trains appeared at 6 of 24 crossings and participants correctly estimated that trains appeared less than half of the time. In the 83% reliability condition, where trains appeared at 5 or 6 crossings (depending on the type of failure), participants responses were similar to second and fourth blocks. For the 50% reliability condition, in which the trains appeared 3 or 6 times, participants correctly responded that trains appeared fewer times than they had in the second and fourth blocks.

In-vehicle Railway Warning Design

At the end of the experiment participants were asked general questions about the IRW system. Nine of the 36 participants said they would always use the IRW system, 17 said often and 9 occasionally. One participant said that he would never use it. While most would use the system, only 14 of 36 (39%) said they would pay for it as an option in a new vehicle. The average acceptable price was $250.00. The amount that people would pay for the system varied from a minimum of $30.00 to a maximum of $1000.00.

Four design questions were asked about the IRW. Thirty-four of 36 participants said that the visual and auditory warning combination expressed an appropriate level of urgency. Thirty-four and 32 also rated the choice of the visual sign and auditory warning appropriate, respectively. Twenty-one (58.3%) said that the warnings came on at the right time in advance of the train crossings. Of the 15 for whom the warnings did not come on at the right time, 14 said the warnings came on too soon. Placement of the IRW was based on the results of Study 1, which indicated that a visual warning should appear about 10 s prior to the railway crossing. The deceleration function may have affected the perception of whether the warnings appeared too soon or at the correct location. At times, when participants pressed the brake after an IRW appeared, then the car/video came to a stop before the train arrived. Participants had to reaccelerate to get closer to the crossings so that the train could come. Some participants translated the need to re-accelerate into the warning coming on too early.

Risky Maneuvers

A number of risk taking measures were used to assess the nature of the sample (see Table 10). Participants were asked if they would perform risky maneuvers on the road if they were running late for important exams at the university. The maneuvers differed in the level of risk. Twenty-four of the 36 participants said that they would never run a red light to get to the university sooner. Nine others said they would do so very rarely, while 3 would run a red occasionally. Half (18 out of 36) would drive at 5-15 kph over the speed limit very often, while another 12 would do it quite often and 5 occasionally. Eighty-nine percent (32 out of 36) would never drive around lowered gates at a railway crossing if the approaching train was clearly some distance away. Two would do so very rarely and another 2 would occasionally do it. It can be argued that responses for this particular risk measure were conservative, given that participants had just finished a study that focused on train warnings. Seventy-two percent of respondents said they either quite often or very often would perform a rolling stop on the way to the university. Another 22% (8) do so occasionally. Participant responses for the remaining questions were more varied. A surprising 31% (11 out of 36) would get angry very often at other drivers for being in the way. Speeding in a playground/school zone and tailgating other drivers were looked down upon by the majority of respondents. No gender differences were found for any of the seven risk taking questions.

The majority of responses reflect that the sample adequately represented the younger population (Rabinovich, 1996). Most respondents would speed 5 to 15 kph over the speed limit if late for an exam. They would also do rolling stops, tailgate, and at times get angry at other drivers. Most would not speed in playground zones, run red lights or drive around lowered gates.

Table 7

Risk taking measure responses

Risk question	very often	quite often	occasionally	very rarely	never
Run a red light	0	0	3 (8.3%)	9 (25%)	24(66.7%)
Drive 5-15 kph over limit	18 (50%)	12 (33.3%)	5 (13.9%)	0	0
Drive around gates	0	0	2 (5.6%)	2 (5.6%)	32(88.9%)
Speed in playground	4 (11.1%)	3 (8.3%)	11 (30.6%)	9 (25%)	9 (25%)
Rolling stop	11 (30.6%)	15 (41.7%)	8 (22.2%)	2 (5.6%)	0
Tailgate other drivers	2 (5.6%)	7 (19.4%)	11 (30.6%)	11 (30.6%)	5(13.9%)
Get angry at other drivers	11 (30.6%)	6 (16.7%)	12 (33.3%)	5 (13.9%)	2 (5.6%)

A composite risk taking score was developed by aggregating participant responses for the seven measures. The range of scores was from 12 to 31, M = 22.4. As the ratings on the seven measures were from 1 to 5, 3 was taken as neutral. Therefore, those with scores above 21 were categorized as high risk takers, and those below 21 were categorized as low risk takers. This resulted in 15 low risk takers and 21 high risk takers. Analyses of the independent measure of risk taking on trust, dependability, or brake press responses were not significant, suggesting that differences in individual risk taking did not differentially effect the results of the study.

Participant Comments

Participant comments were also collected at the completion of the experiment in a post-experiment questionnaire. These are presented in Appendix K, and included comments on the appropriateness of the warning urgency, the visual and auditory warning characteristics, and the timing of the warnings. Generally, the comments were in agreement with the questionnaire responses. The warning was considered urgent enough. The visual sign matched what already appears at railway crossings (i.e., R-AWS). The auditory bell also matched expectancy as it sounds like the bells that already ring at crossings when trains are coming. The timing of the IRW was thought to be right by some and too early by others.

Some participants suggested improving the design of the IRW system so that failures would not occur. They, thus, clearly realized that the failures that they had just experienced had negative consequences. Participants also wanted the visual sign to be more visible (e.g., larger in size) and to better represent the danger with crossing in front of trains (e.g., make it red). They were also concerned about being able to hear the auditory warning in noisy surroundings (e.g., with the stereo on). When approaching crossings, adding the direction of trains' travel was recommended. Some participants wanted to have the capability of modifying the timing of the warnings so that they would come on earlier for when they were driving at higher speeds. Others wanted to keep the warnings on until train hazards had passed.

Participants were also asked hypothetical questions on what they would do if the system failed regularly (i.e., 50% of the time) or infrequently (i.e., 83%). Most responses were similar at the two reliability levels, with responses for lower reliability being more severe compared to those at higher reliability. Most participants would try to get the system fixed, but would reduce their trust in the system. Some would disconnect the system, because of frustration or annoyance (see Lerner et al., 1996 for similar findings).

During false alarms, participants said they would continue to respond to the system, albeit cautiously, because it was better to be safe than sorry. During missed signals, drivers said they would look for other cues (i.e., ringing bells, flashing lights) to inform them about approaching trains. They would rely on their own abilities to detect approaching trains. These comments validate the PRT findings, which showed that even after the false alarms, most participants continued to press the brake.

Study 2: Discussion

The four hypotheses of Study 2 were empirically supported. PRTs to IRW were effected by system failures. Drivers who experienced false alarms braked later. Thus, false alarms did degrade responses to warnings (Lerner et al., 1996). In contrast, drivers who had missed signals braked earlier. These changes in PRTs, for both false alarms and missed signals, were significantly different from baseline, particularly for those in the 50% reliability condition. PRT returned to baseline levels later for those who experienced three false alarms than those who had one false alarm. Driver trust in the system decreased when failures occurred, more so for the 50% reliability condition than the 83% condition. These declines are consistent with previous research (Kantowitz et al., 1997; Muir & Moray, 1996). Trust did return to original levels when the reliability of the system returned to 100% for those in the 83% reliability condition, but not for 50% reliability. Ratings of dependability of the system decreased when failures occurred. For false alarms, decreases in trust were found to be correlated with increases in PRT. Trust was not sensitive to the type of failure that reduced system reliability (see, Kantowitz et al., 1997 for similar findings), whereas the PRT measure was. Subjective comments echoed findings from trust, dependability and PRT analyses. Participants would get annoyed or frustrated if failures occurred regularly (Lerner et al., 1996). If the system failed, then they would take it in for maintenance and would be on the lookout for other cues.

Contextual Factors

The context in which participants interacted with the IRW system clearly placed them in specific circumstances. Participants were provided with the IRW as the only cue about approaching trains. For critical events, the warnings came on and 6 s later the trains arrived, and if participants had pressed the brake during the 6 s, then they were rewarded. As routine safe behaviours are rarely rewarded, the rewards in the study were low. In contrast, costs of unsafe behaviours are high and the penalties in the study were greater than the rewards.

During false alarms, participants were penalized for creating congestion in the traffic behind them. The false alarms placed the participants in no-win situations, and they looked for other cues redundant to IRW that would confirm the warnings. In wanting to avoid congestion and looking for other cues, participants started to brake later. After experiencing false alarms, an alarmingly high number of valid events were simply driven through by participants. Twelve times, 7 participants failed to press the brake (within 6 s) after IRW told them about approaching trains. The missed events occurred because participants did not believe the IRW were valid. The lack of brake press responses clearly replicated instances when repeated false alarms induce persons to not respond at all (e.g., cry wolf folk tale) (see Bliss, Gilson & Deaton for similar findings). Five participants had more than one missed event (5 had two each) and 2 had one each. Each time a missed event took place the participants realized that their lack of response was inappropriate.

During missed signals, participants were charged for crashing. Penalties for crashing were greater than the penalties for congestion, as real costs of crashes are higher than costs of congestion or delay. The missed signals also placed participants in no-win situations, and they started to look for any cues that would tell them that trains were coming. When the sole cue (i.e., the IRW) came on, then participants immediately pressed the brake not wanting to crash. This cautious nature was also exhibited for crossings that didn't have warnings or any trains coming towards them. Therefore, participants in all three failure conditions became more cautious as the result of failures. Those who experienced false alarms did not want to brake prematurely and create congestion in traffic behind them. Participants who experienced missed signals did not want to fail to brake and get in a crash.

Drivers in the False Alarm 1 condition saw and drove through crossings when the false alarms occurred, and thus, were provided with a cue that verified the false alarm. In contrast, drivers in the False Alarm 2 condition received no such cue. While the crossing cue was the only difference between the two false alarm types, it did not result in significant differences between the two false alarm types in PRT, trust or dependability.

External Validity

It is important to discuss the external validity of Study 2 to determine the degree to which the findings generalize to the real driving environments. The purpose of laboratory experiments is not to completely replicate the real world, but to some degree imitate its materials, situations and constraints (Schiff & Arnone, 1995). Participants were provided with dynamic views of approaches to railway crossings that were close to 30 frames per second (fps). The use of summer and winter approaches in the simulation reflect the environmental conditions that effect northern climates. The use of the existing visual and auditory warnings at railway crossings also increased the realism of the experiment. The placement of the visual warning at below the line of horizon slightly above the dashboard on the windshield was done in accord with recommended HUD placements. The sound of the auditory warning emanated from behind the steering wheel which served to direct attention towards the front of the car. Both the visual and the auditory warnings were presented at the same time for the same duration. The auditory warning increased in pitch (i.e., frequency) as the speed of the vehicle increased.

The joining of different driving scenes to simulate a drive to some degree also resembled driving in the real world. When a train appeared, then drivers had to wait 8 s while it moved through the crossing before they could continue on their drive. The simulator was similar to that which is used by drivers (i.e., brake, accelerator and steering wheel), as the controls were bought from a junkyard and removed from real cars. While there was only one deceleration curve (along with the friction coefficient) and three discrete speed levels, they also replicated the functions of devices inside motor vehicles. These devices gave the participants the ability to manipulate the dynamic flow of the video which also resembles driving. The participants were given lower rewards for safe behaviour than penalties for unsafe behaviours. The rewards and penalties placed drivers in a similar mindset as actual driving. While the term congestion described situations during which a driver unnecessarily slows down or stops, the use of it, as a negative outcome, includes a driver's own perception of his/her delay. The use of these materials for the study assisted participants in quickly understanding how to use the system.

While field research is needed to validate Study 2 findings, they generalize to railway crossings that have warnings which may give false alarms or missed signals (e.g., flashing lights, warning bells), especially if these crossings are prone to poor visibility (e.g., fog, vegetation, buildings). Beyond warnings for trains, these findings can also be applied to false alarms and missed signals for other hazard situations. For example, false alarms at construction sites may reduce driver trust of the warnings, coupled with a need to look for redundant cues to confirm the warnings. After experiencing false alarms drivers may not start slowing down or may even drive by until the hazard is confirmed through another cue. This would especially be so if previous responses to false alarms had cost drivers in some way. On the other hand, missed signals may make drivers approach construction sites more cautiously by slowing their vehicles.

The findings of Study 2 are based on both objective and subjective measures. Participants did consider their interaction with the system as important. They were aware of the costs and implications of being in crashes. They did not drive at high speeds and did not treat the system as if it were a video game. Some participants did fail to press the brake to more than one critical event, but realized that their lack of response was inappropriate. The verbal protocol made explicit the importance of the research topic, and all participants treated the reward/penalty structure as real-life like. Subjective comments do reveal some conscientious participant suggestions and recommendations for improving IRW. The simulation did, however, place drivers in a somewhat contrived situation and field experimentation is required to validate Study 2 findings.

Finally, while two reliability levels (83% and 50%) were used in the study, and it can be argued that no warning system will be 50% reliable, it was the third block that was 83% or 50% reliable. In other words, changes in PRT, trust and dependability present in the 83% reliability condition were due to a single failure (either false alarm or missed signal). Similarly, the larger changes in the 50% reliability condition were because of three failures within short time periods. Therefore, the differential reliability results do generalize to IRW systems that may have a single failure (as 83%) or many failures (as 50%).

Limitations of Present Study

All four possibilities in the states-of-the-world matrix of signal detection were included in Study 2. The possibilities were: train present/warning present (i.e., valid critical events), train absent/warning present (i.e., false alarms), train present/warning absent (i.e., missed signals), train absent/warning absent (i.e., events). The inclusion of the missed signal is an expansion of the design of Lerner et al. (1996) and their test of annoyance with excessive false alarms. The results have shown that false alarms and missed signals have differential effects on how drivers respond to subsequent valid events. The use of failure type as a between-participant variable, however, limited all participants from experiencing each failure and being able to compare between them. Conversely, Study 2 methodology eliminated the possibility of the confound of failure type, which would have been present had each participant experienced more than one type of failure. Future research could give participants experience with all types of failures and measure both subjective and objective responses.

Study 2 imposed a single task on the participants, that being to drive the car safely. While this included braking and accelerating, the steering wheel control was not functional. Similarly, other aspects of driving (e.g., stopping at a red light or merging onto a highway) were not included in the experimental design, that may have made the task relatively easy to perform. The use of the reward/penalty structure did, however, impose cognitive demand on drivers similar to real driving situations.

The issue of whether the short term novelty effects (e.g., third and fourth blocks in Study 2) that were found are generalizable to long term habit forming effects (e.g., many years) still remains. It has been shown that performance and trust with a system revert back to original levels after failures have ceased to occur. Does continual experience with faulty systems permanently effect both subjective trust and objective performance? While research has shown that it does within the time and location constraints of a laboratory experiment (Lee & Moray, 1994; Muir & Moray, 1996), further research in the real world using innovative methodologies is required.

Areas for Future Research

While the study was performed with younger drivers, other groups of individuals would also benefit from in-vehicle warnings at railway crossings. For example, bus drivers are placed in situations where they not only have to drive safely, but also have to monitor students on their buses. There is the possibility that, at times, bus drivers may become distracted by persons within their bus and not notice cues in the external environment that identify approaching trains. Bus drivers would benefit from an IRW, especially if it was in both visual and auditory modes. Truck drivers, who travel on long hauls, and suffer from low arousal and tiredness, would also benefit.

The visual IRW was a replication of the existing sign warning post-mounted in advance of railway crossings. This sign (R-AWS) depicts a roadway intersecting a set of railway tracks. Future research could validate the choice of the visual warning. Different types of visual warnings may also be tested to warn drivers about different occurrences at railway crossings (e.g., trains from the left or right, or about the second train phenomenon). While the study assumed that placement of the in-vehicle warning as a HUD would not be problematic, as it gives primary hazard information to the driver, there are a number of unresolved HUD issues (such as Gish & Staplin, 1995). As more hazard information is deemed presentable in a HUD, the issues of visual space clutter and the order of presentation (i.e., which information is presented when) will need to be addressed.

Application of In-vehicle Warnings in HUDs

Gish and Staplin (1995) provided a review of perceptual and cognitive issues with using HUDs in automobiles. High priority perceptual factors include spatial location and luminance contrast. Images in the HUD should not block or mask objects in the external environment. At low background luminance, there is a potential for the luminance contrast of the image in the HUD to be too high, and vice versa. Luminance contrast is the difference between the luminance of an object in the foreground and the luminance of the background environment. There is no optimal level of contrast for all driving conditions, and the luminance of a HUD must be adjustable. The amount and format of information in HUDs are important cognitive issues. Excess information can lead to cognitive and attentional capture, distracting drivers from noticing other important objects on roadways (e.g., pedestrians). In moving beyond simple head-up displays of vehicle speed into more critical information (e.g., advisories of road conditions, warnings about approaching hazards), "reliable measures of the effect of HUD use on responses to priority external targets must be obtained, under realistic operating conditions" (Gish & Staplin, 1995, p. xiii). Grant, Kiefer and Wierwille (1995) report that there is a greater probability of detecting and identifying briefly (transitory) presented information in a head-up format (i.e., on the windshield) than in a head-down format (i.e., below the dashboard). While the probability of detection/identification is increased, there is some doubt on the type of response that is elicited when a driver sees an image displayed in the HUD. For example, if the display of transitory telltale warnings is incorporated into the HUD and the word ENGINE appears on the windshield, what does this display tell the driver, what should it tell the driver and what is the appropriate response? The test of in-vehicle warnings in real HUDs is required, under more naturalistic conditions with a wide variety of both independent and dependent variables, to address the aforementioned issues.

Implications for Design

Study 2 results have raised some critical practical design issues for IRW. Signal detection theory concepts can be applied to the issues (Swets, 1992). Sensitivity (d') of the in-vehicle warning system is determined by the ability of the sensors to separate noise from the signals. Clearly, sensitivity should be high. This would result in high reliability, and Study 2 has shown that trust in the system remains high if the system is highly reliable. More importantly, high reliability allows drivers to develop and practice accurate expectancies. The criterion (beta) is presumably set by designers to achieve different hit/false alarm rates for a system (Hancock et al., 1995). Study 2 has shown that participants respond differently after experiencing false alarms than after missed signals. For those crossings that are already actively controlled (i.e., they have bells, lights, gates), it would be beneficial to designers to choose a criterion that will have some missed signals in wanting to reduce false alarms. Participants have shown that after experiencing missed signals, they get more cautious and look for any cues that tell them that a train is coming. If missed signals do occur, then there are redundant cues available (e.g., bells, lights).

For those crossings that are not actively controlled then the criterion should allow some false alarms in wanting to reduce missed signals. For these crossings there are no other active cues available to the driver (except for the train) and any missed signals will not fulfill the purpose of an in-vehicle warning system. Farber and Paley (1993) suggested that some false alarms may also induce the operator to increase vigilance. While brake press responses were shown to be effected by false alarms, the increase would still give the driver enough time to stop the vehicle before it reached the crossing (especially if the onset of warning is at 10 s). The 12 missed events are distressing, but it is the hope that at least drivers will look around for a train even after experiencing false alarms.

Finally, the IRW as designed for Study 2 appears to be an acceptable one. The visual aspect was a replication of the existing external signs, and benefits include established expectancy and iconic nature. The auditory warning was also a replication of the existing sound heard at railway crossings, and it similarly benefits from established expectancy, along with it being acoustic rather than a verbal. Some participants mentioned that the warning should be at a volume that can be heard over existing sounds inside the car (e.g., radio, conversations). The warning onset time of 2.5 s was determined by previous research and was long enough to mandate attention but not long enough to annoy.

General Discussion

From 1991 to 1994, 71% of fatal railway crossing accidents were the result of either driver disobedience of warning signals or failing to yield right-of-way to trains (TC, 1996). Åberg (1988) and Knoblauch et al. (1982) have stated that, at times, drivers can be distracted and do not notice/recognize externally placed warnings. Drivers also intentionally make ill-fated judgments in deciding to violate warnings, due to some degree of motorist uncertainty of when the train will actually arrive at the crossing, relative to when the bells and/or light warnings are activated (Leibowitz, 1985).

The results of Study 1 have shown that drivers consistently chose 10 s as advance warning time for an approaching train before they reach the crossings. Future research will, however, need to address how long before a train arrives at a crossing should in-vehicle railway warnings be activated. The findings of Study 2 have shown that, after experiencing false alarms, drivers try to second guess warnings, and behaviour becomes less predictable and more variable. Drivers seem to be searching for redundant cues that will confirm the IRW. While missed signals have costs at instances when they occur, driver behaviour becomes more cautious in responding to events after experiencing a number of missed signals. High levels of reliability naturally result in high levels of driver trust in IRW. High reliability is important as it cues the driver to initiate a response (i.e., start decelerating by either braking or letting off the accelerator). Once a response is initiated it gives drivers more time to comfortably and safely complete the chosen maneuver.

In-vehicle railway warnings are not the only human factors solutions in pursuit of safe traffic systems at railway crossings. Clearly, a bridge which separates train traffic from vehicle traffic is the ultimate engineering intervention. Another engineering intervention is the removal of train tracks from certain crossings by re-routing train flow through other tracks. This reduces the number of crossings, but would also increase train traffic flow at other crossings. Accident data have shown that more accidents occur at crossing with more train traffic. Other engineering interventions include placing traffic islands in the middle of roadways to deter drivers from driving around lowered gates (and onto opposing traffic lanes). Educational programs targeting specific types of drivers who intentionally violate railway warnings may also be used. Finally, rather than focusing on enforcement of warnings which may not be practiced (see Abraham et al., 1998), in-vehicle warnings can be used to encourage drivers to comply with warnings. Extensive research focusing specifically on what type of information can be given to drivers is needed so that it encourages them to practice safe behaviour rather than induces them to commit further unsafe acts. While the use of IRW do not in themselves suggest that drivers will cease to violate train right-of-ways, their testing, development, and evaluation are important first steps in rethinking the interaction between trains and vehicles so that violations do not occur or are discouraged because of system design.

References

Åberg, L. (1988). Driver behaviour at flashing-light, rail-highway crossings. Accident Analysis & Pervention, 20(1), 59-65.

Abraham, J., Datta, T. K. & Datta, S. (1998). Driver Behaviour at Rail-Highway Crossings. Presented at the 77th Annual Meeting of the Transportation Research Board. Washington, DC.

Alberta Transportation and Utilities. (1997a). Operator's Licence Information: All Classes Including Cars and Light Trucks (Basic Licence Driver's Handbook). Edmonton, AB: Alberta Transportation and Utilities.

Alberta Transportation and Utilities. (1997b). Operator's Licence Information: Large Trucks, Tractor Trailers, Ambulances, Taxis and Buses (Supplement). Edmonton, AB: Alberta Transportation and Utilities.

Alexander, G. J. (1991). Search and perception-reaction time at intersections and railroad grade crossings. ITE Journal, 59(11), 17-20.

Alexander, G. J., & Lunenfeld, H. (1975). Positive Guidance in Traffic Control. Washington, DC: US Dept. of Transportation, Federal Highway Administration.

Alexander, G. J., & Lunenfeld, H. (1984). A users' guide to positive guidance in highway control. In R. Easterby & H. Zwaga (Eds.), Information Design (pp. 351-383). New York : John Wiley & Sons.

Bliss, J. P., Gilson, R. D., & Deaton, J. E. (1995). Human probability matching behaviour in response to alarms of varying reliability. Ergonomics, 38(11), 2300-2312.

Caird, J. K., & Chugh, J. S. (1997). The time cost of head-up displays for older drivers: Critical event onset, task location, and display type. Proceedings of the 41st Annual Conference of the Human Factors and Ergonomics Society (pp. 1008 - 1012). Santa Monica, CA: Human Factors and Ergonomics Society.

Carroll, A. A., & Helser, J. L. (1996). Safety of Highway-Railroad Grade Crossings Research Needs Workshop, 1 (DOT-VNTSC-FRA-95-12.1). Washington, DC: US Dept. of Transportation, Federal Highway Administration.

Cerrelli, E. (1994). Female drivers in fatal crashes: Recent trends (Rep. No. DOT/HS 808 106). Washington, DC: National Highway Traffic Safety Administration.

Chugh, J. S., & Hove, P. (1997). A pilot study of advance in-vehicle warnings for railway crossings. (Report No. 97-01). Calgary, AB: Cognitive Ergonomics Research Laboratory, University of Calgary.

Cohen, J. (1977). Statistical power analysis for the behavioural sciences (Rev. Ed.). New York: Academic Press.

Council of Uniform Traffic Control Devices for Canada (1976). Uniform Traffic Control Devices for Canada. Ottawa, ON.

Dewar, R. (1988). Criteria for the design and evaluation of traffic sign symbols. Transportation Research Record, 1160 (pp. 1-6). Washington, DC: Transportation Research Board.

Dewar, R. E. (July 1993). Warning: Hazardous road signs ahead. Ergonomics In Design, 26-31.

Dewar, R. E., & Ells, J. G. (1984). Methods of evaluation of traffic signs. In R. Easterby & H. Zwaga (Eds.), Information Design. (pp. 77-90). New York: John Wiley & Sons.

Dewar, R. E., Ells, J. G., & Mundy, G. (1976). Reaction time as an index of traffic sign perception. Human Factors, 18(4), 381-392.

Ells, J. G., Dewar, R. E., & Milloy, D. G. (1990). An evaluation of six configurations of the railway crossbuck sign. Ergonomics, 23, 359-367.

Fambro, D. B., Koppa, R. J., Picha, D. L., & Fitzpatrick, K. (1998). Driver perception-brake response in stopping sight distance situations. Transportation Research Record, 1628 (pp. 1-7). Washington, DC: Transportation Research Board.

Farber, E., & Palley, M. (1993). Using freeway traffic data to estimate the effectiveness of rear end collision countermeasures. Paper presented at the Third Annual IVHS America Meeting, Washington, DC.

Federal Railroad Administration. (1996). Highway-rail crossing accident / incident and inventory bulletin: Calendar year 1995. (Report No. 18). Washington, DC: US Dept. of Transportation, Federal Railroad Administration.

Fisher, J. (1992). Testing the effect of road traffic signs' informational value on driver behaviour. Human Factors, 34(2), 231-237.

Gish, K. W., & Staplin, L. (1995). Human factors aspects of using head-up displays in automobiles. (Interim Report DOT HS 808 320). Washington, DC: US Dept. of Transportation, Federal Highway Administration.

Glennon, J. C. (1995). Rail-highway grade crossing safety. In J. R. Loumiet & W. G. Jungbauer (Eds.), Train accident reconstruction and FELA & railroad litigation (2nd ed.). Tuscon, AZ: Lawyers & Judges.

Grant, B. S., Kiefer, R. J. & Wierwille, W. W. (1995). Drivers' detection and identification of head-up versus head-down telltale warnings in automobiles. Proceedings of the 39th Annual Conference of the Human Factors and Ergonomics Society (pp. 1087 - 1091). Santa Monica, CA: Human Factors and Ergonomics Society.

Haas, E., & Casali, J. G. (1995). Perceived urgency and response time to multi-tone and frequency-modulated warning signals in broadband noise. Ergonomics, 38(11), 2313-2326.

Hancock, P. A., Parasuraman, R., & Byrne, E. A. (1995). Driver-centered issues in advanced automation for motor vehicles. In R. Parasuraman, & M. Mouloua. (Eds.), Automation and Human Performance: Theory and Applications. (pp. 337-364). Hillsdale, NJ: Erlbaum.

Hawkins, H. G. (1994). Use of supplemental plaques to improve effectiveness of warning signs. Transportation Research Record, 1456 (pp. 20-26). Washington, DC: Transportation Research Board.

Hochberg, J., & Brooks, V. (1978). Film cutting and visual momentum. In J. W. Senders, D. F. Fisher, & R. A. Monty (Eds.), Eye Movements and the Higher Psychological Functions. (pp. 293-313). Hillsdale, NJ: Erlbaum.

Huey, R. W., Harpster, J. L., & Lerner, N. D. (1997). Summary report on in-vehicle crash avoidance warning systems: Human factors considerations. (Rep. No. DOT HS 808 531). Washington, DC: US Dept. of Transportation, Federal Highway Administration.

Kaiser, M. K., & Phatak, A. V. (1993). Things that go bump in the light: On the optical specification of contact severity. Journal of Experimental Psychology: Human Perception and Performance, 19(1), 194-202.

Kantowitz, B.H. , Hanowski, R. J., & Kantowitz, S. C. (1997). Driver acceptance of unreliable traffic information in familiar and unfamiliar settings. Human Factors, 39(2), 164-176.

Killick, A. N., Pineau, N. S., Moodie, E. R., DesJardins, C. M., & Caird, J. K. (1997). The effect of in-vehicle signs and traffic context on driver's comprehension and reaction time. Proceedings of the 29th Annual Conference of the Human Factors Association of Canada (pp. 59-67). Mississauga, ON: Human Factors Association of Canada.

Klein, T., Morgan, T., & Weiner, A. (1994). Rail-highway crossing safety fatal crash and demographic descriptors (Pub. No. DOT HS 808 196). Washington, DC: US Department of Transportation, National Highway Traffic Safety Administration.

Kline, D. W., & Fuchs, P. (1993). The visibility of symbolic highway signs can be increased among drivers of all ages. Human Factors, 35(1), 25-34.

Kline, T. J. B., Ghali, L. M., Kline, D. W., & Brown, S. (1990). Visibility distance of highway signs among young, middle-aged, and older observers: Icons are better than text. Human Factors, 32(5), 609-619.

Knoblauch, K., Hucke, W. & Berg, W. (1982). Rail Highway Crossing Accident Causation Study Volume I. Executive Summary (Pub. No. DOT-FH-11-9682). Washington, DC: US Department of Transportation, Federal Highway Administration.

Lee, J., & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. Ergonomics, 35(10), 1243-1270.

Lee, J., & Moray, N. (1994). Trust, self-confidence, and operators' adaptation to automation. International Journal of Human-Computer Studies, 40, 153-184.

Leibowitz, H. W. (1985). Grade crossing accidents and human factors engineering. American Scientist, 73, 558–562.

Lerner, N., Ratte, D., & Walker, J. (1990). Driver behaviour at rail-highway crossings (Rep. No. FHWA-SA-90-008). Washington, DC: US Dept. of Transportation, Federal Highway Administration.

Lerner, N. D., Dekker, D. K., Steinberg, G.V., & Huey, R.W. (1996). Inappropriate alarm rates and driver annoyance (Rep. No. DOT HS 808 533). Washington, DC: US Dept. of Transportation, Federal Highway Administration.

Meeker, F., Fox. D., & Weber, C. (1997). A comparison of driver behaviour at railroad grade crossings with two different protection systems. Accident Analysis and Pervention, 29(1), 11-16.

Muir, B. (1994). Trust in automation: Part I. Theoretical issues in the study of trust and human intervention in automated systems. Ergonomics, 37(11), 1905-1922.

Muir, B., & Moray, N. (1996). Trust in automation: Part II. Experimental studies of trust and human intervention in a process control simulation. Ergonomics, 39(3), 429-460.

Nelson, C. (1998). [Sixteen train-related accident descriptions]. Unpublished raw data. City of Calgary, AB.

Nohre, L., MacKinnon, D.P., & Sadalla, E.K. (1998). The effects of warning false alarms on driving speed. Poster presentation at the 42nd Annual Conference of the Human Factors and Ergonomics Society (p. 1627). Santa Monica, CA: Human Factors and Ergonomics Society.

Olson, P. L. (1996). Forensic aspects of driver perception and response. Tuscon, AZ: Lawyers & Judges.

Olson, P. L., & Sivak, M. (1986). Perception-response time to unexpected roadway hazards. Human Factors, 28(1), 91-96.

Operation Lifesaver. (1997). Highway-rail facts & terms. [On-line]. Available : http://www.ncol.org

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. Human Factors, 39(2), 230-253.

Poulton, E. C. (1982). Influential companions: Effects of one strategy on another in the within-subjects designs of cognitive psychology. Psychological Bulletin, 91(3), 673-690.

Rabinovich, B. A. (1996). Innovative methods for training to improve young driver risk perception. (Rep. No. NSF-9561298). Arlington, VA: National Science Foundation.

Rasmussen, J. (1983). Skills, rules, and knowledge: Signals, signs, and symbols, and other distinctions in human performance models. IEEE Transactions on Systems, Man, and Cybernetics, 13(3), 257-266.

Rempel, J. K., Holmes, J. G., & Zanna, M. P. (1985). Trust in close relationships. Journal of Personality and Social Psychology, 49, 95-112.

Richards, H. A., & Bartoskewitz, R. T. (1995). The intelligent highway-rail intersection integrating ITS and ATCS for improved grade crossing operation and safety (Draft Rep. UMTRI-89436, pp. G-35-G52). Washington, DC: US Dept. of Transportation, Federal Highway Administration.

Richards, S. H., & Heatington, K. W. (1988). Motorists' understanding of railroad-highway grade crossing traffic control devices and associated traffic laws. Paper presented at the 67th Annual Meeting of the Transportation Research Board. Washington, DC.

Road Safety Directorate. (1995). Road/railway grade crossing manual (Draft). Ottawa, ON: Railway Safety Directorate, Surface Group, Transport Canada.

Schieber, F. (1995). Fourier techniques for optimizing symbol highway sign legibility. Presented at the 65th ITE Annual Meeting. Denver, CO.

Schiff, W., & Arnone, W. (1995). Perceiving and driving: Where parallel roads meet. In P. Hancock, J. Flach, J. Caird & K. Vicente (Eds.), Local Applications of the Ecological Approach to Human-Machine Systems (pp. 1-35). Hillsdale, NJ: Erlbaum.

Standards Australia. (1991). Manual of uniform traffic control devices. North Sydney, Australia.

Staplin, L., & Fisk, A. D. (1991). A cognitive engineering approach to improving signalized left turn intersections. Human Factors, 33(5), 559-571.

Swets, J. A. (1992). The science of choosing the right decision threshold in high-stakes diagnostics. American Psychologist, 47(4), 522-532.

Tan, A. K., & Lerner, N. D. (1995). Multiple attribute evaluation of auditory warning signals for in-vehicle crash avoidance warning systems. (Rep. No. DOT HS 808 535). Washington, DC: US Dept. of Transportation, Federal Highway Administration.

Tardiff, L., Parviainen, J., & Ede, W. J. M. (1996). Application of intelligent transportation systems (ITS) / advanced train control systems (ATCS) technologies at highway-rail level crossings. (Res. Rep.). Ottawa, ON: Transport Association of Canada.

Transport Canada. (1998). Safety railway accidents. [On-line]. Available : http://www.tc.gc.ca/tfacts/StatFor/safety/safety6.htm

Transport Canada. (1996). Railway safety facts 1996: Crossing accidents. Ottawa, ON: Transport Canada.

Transportation Safety Board. (1997a). Forward. [On-line]. Available : http://bst-tsb.gc.ca/stats/e1996rail1.html

Transportation Safety Board. (1997b). Table 1 : Railway occurrences and causalities 1987 - 1996. [On-line]. Available : http://bst-tsb.gc.ca/stats/e1996rail2.html

Transportation Safety Board. (1997c). Table 2 : Fatalities / injuries by type of occurrence and person type 1987 - 1996. [On-line]. Available : http://bst-tsb.gc.ca/stats/e1996rail3.html

Transportation Safety Board. (1997d). Table 6 : Crossing accidents and causalities by type of crossing and protection 1987 - 1996. [On-line]. Available : http://bst-tsb.gc.ca/stats/e1996rail7.html

Tufano, D. R., Knee, H. E. & Spelt, P. F. (1996). In-vehicle signing functions of an in-vehicle information system. Proceedings of the Intelligent Transportation Systems (pp. 182-188). Houston, TX: ITS America.

Uniform Traffic Control Devices for Canada. (1976). Ottawa, ON: Roads & Transportation Association of Canada.

Ward, N. J., & Wilde, G. J. S. (1995a). A comparison of vehicular approach speed and braking between day and nighttime periods at an automated railway crossing. Safety Science, 19, 31-44.

Ward, N. J., & Wilde, G. J. S. (1995b). Field observation of advance warning/advisory signage for passive railway crossings with restricted lateral sightline visibility: An experimental investigation. Accident Analysis and Prevention, 27(2), 185-197.

Ward, N. J., & Wilde, G. J. S. (1996). Driver approach behaviour at an unprotected railway crossing before and after enhancement of lateral sight distances: An experimental investigation of a risk perception and behavioural compensation hypothesis. Safety Science, 22, (1-3), 63-75.

Zaidel, D. M. (1991). Specification of a Methodology for Investigating the Human Factors of Advanced Driver Information Systems (Pub. No. TP 11199). Ottawa, ON: Transport Canada, Ergonomics Division, Road Safety and Motor Vehicle Regulation.

Appendix A
Pre-experiment questionnaire: Study 1

Participant #_______

Age ________years

Gender M / F

Driving Experience _______years

Km driven per year _________

Is the majority of your driving in: Rural / Urban areas

Are you familiar with the use of a VCR remote control? yes / no

Have you driven a car with a head-up display device? yes / no

Do you wear glasses or contact lenses that correct your vision? yes / no

<u>For experimenter testing:</u>

Visual acuity 20/____

Contrast sensitivity A___ B___ C___ D___ E___

Appendix B
Experimental protocol: Study 1

Thank you very much for coming here to participate in this experiment. You are free to withdraw from the study at any time.

Are you comfortable in your seating; we can adjust your chair by raising it or lowering it if you like. RAISE OR LOWER IF NECESSARY. We would like to keep the distance from the TV as it is, because it simulates the distance of the drivers eyes from the windshield where head up displays are shown.

Head up displays are digital images that can be displayed on the windshield of a car. These images can include speedometer information, turn signals, among others. Today we want you to interact with this video recorder in selecting locations where railway warning signs would appear in a HUD.

To do this, I will show you a video clip that depicts a driver's approach to a railway crossing. Then I will get you to use this remote control, and by rewinding and forwarding through the clip, I want you to select the location when an advance warning sign would be displayed on the windshield. These clips were recorded during both winter and summer seasons. As the camera was placed near the driver's seat, assume that you are the driver of the car for this test.

You should pick a location that you feel will give you enough time to notice that the HUD sign is there, to look around for an approaching train, and if need be to stop the car before reaching the crossing.

Do you have any questions about what I would like you to do?

ANSWER QUESTIONS.

Are you ready for some practice with the system.

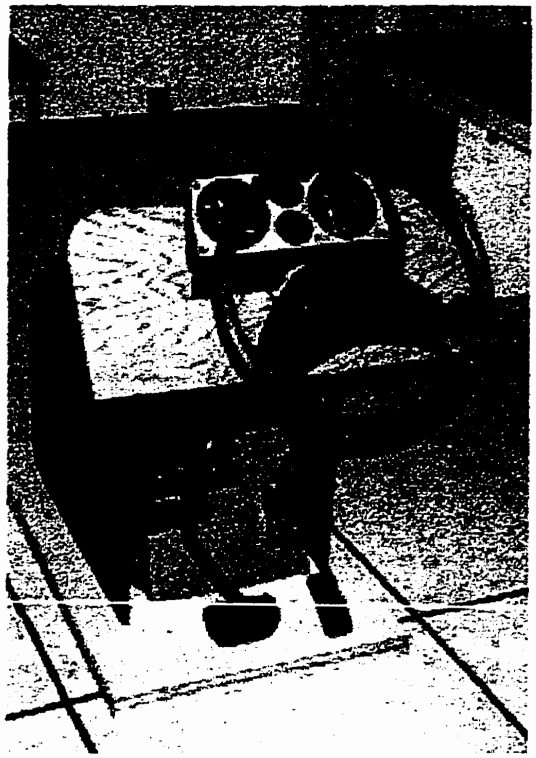
PRACTICE WITH THE FOUR EXAMPLE CLIPS.

Was that O.K.? After seeing a clip and selecting a location, you are more than welcome to look somewhere else to rest your eyes. If at any time you feel that your eyes are getting tired or you want to take a break, just let me know. Okay.

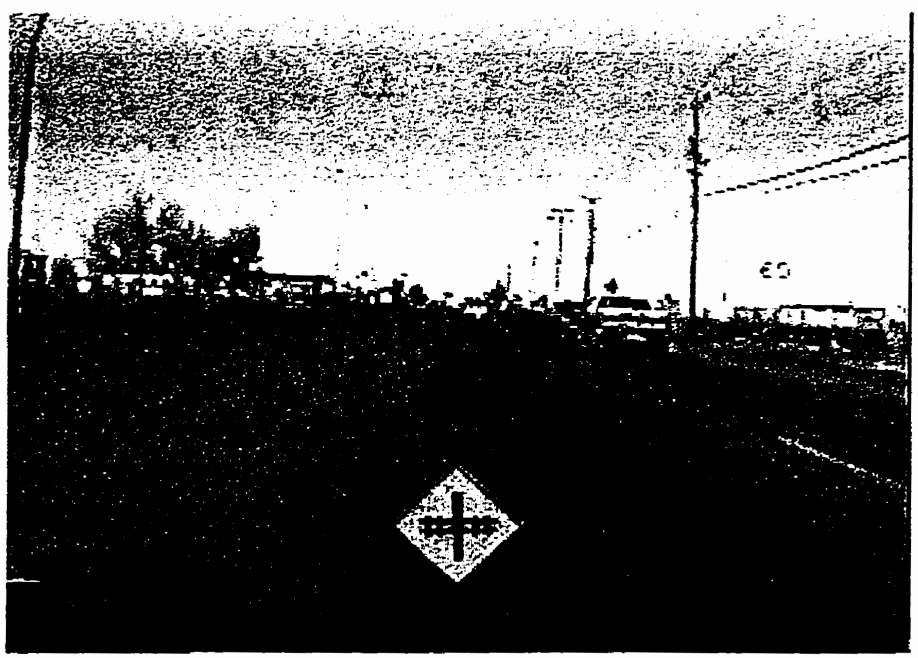
Should we start the experimental clips.

SIXTEEN EXPERIMENTAL CLIPS; OFFER BREAK IF REQUESTED OR DEEMED.

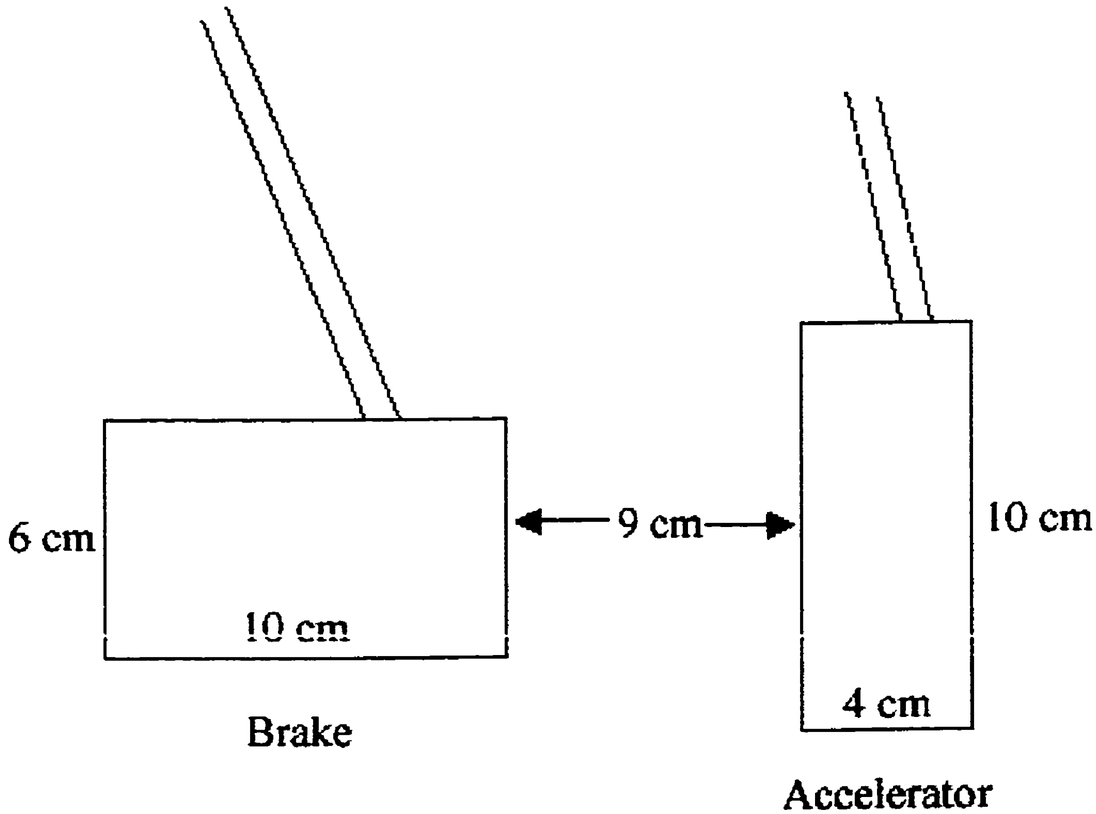
Appendix C
Low-fidelity simulator, IRW, and brake/accelerator schematic



Low-fidelity simulator



In-vehicle Railway Warning (IRW)



Brake and accelerator schematic

Appendix D
Pre-experiment questionnaire: Study 2

Participant #______

1. Age ________years
2. Gender M / F
3. Driving Experience _______years
4. Km driven per year _________ (20,000 is average)
5. Is the majority of your driving in: Rural / Urban areas
6. Have you driven a car with a head-up display device? yes / no
7. Do you wear glasses or contact lenses that correct your vision? yes / no
8. How many hours in a week do you use a computer? _____ hrs

In this experiment, we are interested in your judgments of trust in the train warning system. First, please think about your trust in people. We do not place equal amounts of trust in all the people that we know, and we can express how much trust we place in a person. We can also place trust in products (e.g., cars and computers). For example, I trust my Honda to start in the morning because it has never failed to do so. In contrast to our trust that a machine can perform a purpose reliably, we can also rate our confidence in ourselves to perform certain activities. For example, a person who is not afraid of heights and has climbed trees when they were younger may have a relatively high degree of self confidence in being able to climb a ladder to paint the roof.

Trust and self-confidence can be expressed on scales similar to the ones below. Remember there are no right or wrong answers. To gain some practice with the trust scale, please rate your trust in each of the following:

1) How much do you trust your calculator to give you the correct answer once the numbers have been entered. The calculations are related to trigonometry. Circle score.

1	2	3	4	5	6	7
absolutely no trust						absolute trust

How confident are you in your ability to manually calculate the same operations.

1	2	3	4	5	6	7
absolutely no confidence						absolute confidence

How confident are you in your knowledge of how to effectively use a scientific calculator (i.e., know which operations to enter).

1	2	3	4	5	6	7
absolutely no confidence						absolute confidence

2) How much do you trust your watch to tell you the correct time.

1	2	3	4	5	6	7
absolutely no trust						absolute trust

How confident are you in your ability to estimate the correct time by looking at the sun's position.

1	2	3	4	5	6	7
absolutely no confidence						absolute confidence

For experimenter testing:

Visual acuity 20/____
Contrast sensitivity A___ B___ C___ D___ E___

Appendix E
Post-block questionnaire: Study 2

Participant #_____

Please answer the following questions. Again, remember there are no right or wrong answers. Your answers will help us in our search of how to improve the relation between the driver and the advance warning system.

1) How often did a train appear within each trial?

1	2	3	4	5	6	7
never	rarely	less than half	half of the time	more than half	more often than not	always

2) How much do you trust the in-vehicle warning system to provide you with advance warning about trains?

1	2	3	4	5	6	7
absolutely no trust						absolute trust

3) How dependable do you think the in-vehicle warning system is?

1	2	3	4	5	6	7
absolutely no dependability						absolute dependability

Appendix F
Post-experiment questionnaire: Study 2

Participant #_______

1) Do you think that the auditory and visual warnings expressed an appropriate level of urgency?

   Yes / No

   Why or Why Not?

2) Was the choice of the visual sign appropriate?

   Yes / No

   Why or why not?

3) Was the choice of the auditory warning appropriate?

   Yes / No

   Why or why not?

4) Did the in-vehicle warnings come on at the right time (i.e., not too soon/not too late)?

   Yes / No

   Why or why not?

5) Can you suggest design improvements to the in-vehicle warning system?

6) How much do you now trust the warning system to provide you with advance warning about trains?

1	2	3	4	5	6	7
absolutely no trust						absolute trust

7) Was there ever a time when the warning system did not appear to work?

yes / no

8) If so, describe what happened.

9) What would you do if the failures occurred on a regular basis (i.e., half of the time)?

10) on an infrequent basis (i.e., 10-20 % of the time)?

11) If you had this in-vehicle warning system operating in your car, how often would you use its information?

1	2	3	4	5
never	rarely	occasionally	often	always

12) Would you pay extra to get it installed as an option in a new car?

Yes / No If yes, how many dollars ___________

13) How confident are you in your ability to install the new Microsoft Windows 98 application onto a hard drive of a PC?

1	2	3	4	5	6	7
absolutely no confidence						absolute confidence

14) How confident are you in your ability to merge onto a high speed highway (e.g., Deerfoot Trail) during high congestion times?

1	2	3	4	5	6	7
absolutely no confidence						absolute confidence

15) If you were in a hurry to get to an important exam for a class how often would you (remember there are no right or wrong answers):

1: Very often	2: Quite often	3: Occasionally	4: Very rarely	5: Never

a) Run a red light to get to the university sooner ___
b) Drive at 5-15 kph over the speed limit ___
c) Drive around lowered gates at a railway crossing if the approaching train was clearly some distance away. ___
d) Speed in a playground / school zone ___
e- Do a rolling stop (i.e., not a complete stop) at a stop sign ___
f) Tailgate other people to get them to drive faster ___
g) Get angry at other drivers for being in your way ___

Thank you for participating in this experiment !!

Appendix G
Experiment Protocol: Study 2

Thank you very much for coming here to participate in this experiment. You are free to withdraw from the study at any time.

Please take a seat in front of this apparatus. We can bring it closer or move it further from you if you like. You can also raise and lower the chair to suit your comfort.

Today I will show you some video segments of a motorist driving through the city. These video segments were taken from inside the vehicle looking through the windshield. You will exactly what a driver would see. Approximately half of the segments were recorded during the summer and the other half were during the winter. The video scenes have been edited together to simulate a normal drive.

You can use the brake and the accelerator to control the rate at which the video segments play. When you press the accelerator the video will start at normal speed. Press it a little more and the speed rises to one and a quarter the speed, and if you press it all the way then the speed is one and a half times. This simulates real driving to some degree.

Pressing the brake reduces the speed at which the movies are played. If you keep the brake pressed the video will come to a stop. To restart the video simply press the accelerator. If you release the accelerator and don't press on the brake then the video will slowly come to a stop.

You can rest your hands/arms on the steering wheel, and move it around during the test. Moving the steering wheel does not however change the information on the screen. When you go around curves you should turn the steering wheel in the direction of the turn.

- SHOW BRAKE CLIP
- PRACTICE WITH BRAKE / ACCELERATOR
- MENTION WEATHER SUMMER/WINTER

Some of the driving scenes go through railway crossings and others do not. When you approach a railway crossing, there can be two possible scenarios. One, there is no train coming towards the crossing. In this case, there is no need to stop the vehicle, and as in real driving you would drive through.

In the second case, a train is coming towards the crossing. During these instances the safest strategy would be to wait for the train, and then go through the crossing when the coast is clear (i.e., no train coming from the other direction).

A new in-vehicle warning system has been developed that detects an oncoming train and sends a signal to your car. The car then displays a warning sign on the windshield. An auditory warning supplements (i.e., comes on at the same time) the visual warning. These warnings are designed to tell the driver that a train is coming.

Today, we are going to test the effectiveness of this in-vehicle warning system.

While watching the video segments, these warnings may come on. When they do they are telling you that a train is coming to the crossing. Practicing safe behaviour you should bring your vehicle to a stop before it reaches the crossing, and before a train comes.

Once you have stopped at a crossing, and when the train has gone by, you can move your foot off the brake and onto the accelerator to restart the video scenes. In this experimental setup no trains will be coming from the other side so you don't have to worry about the second train phenomenon.

If the sign comes on and you press the brake and the car stops and nothing happens then just press the accelerator a little bit. During these instances you have stopped the car a little too far away from the crossing, and you need to bring it a little closer to the crossing for the train clip to be activated. Make sure that you wait for the train to pass through and keep your foot on the brake when the train is present.

- SHOW WARNING CLIPS
- EXPLAIN EVENTS (warning sign/bells)

As we are trying to make this simulation as life like as possible we have included a payoff matrix. Let me explain. When we do something beneficial to our survival we are rewarded and when we do something that poses risks to our survival then at times we are penalized.

Therefore, when you brake to slow down the car to bring it to a stop and a train comes then you will be rewarded with 50 cents. This represents safe behaviour.

When you brake and your car stops and there is no train coming, then this will cost you $1.00. In this case you have braked for no hazard and created congestion in the traffic behind you. Reflecting the cost of $1.00.

If you fail to brake for a train, then this will cost you $2.00. This clearly represents instances when a collision between a train and a vehicle can occur.

This payoff matrix was designed to be somewhat real-life like, as routine safe behaviours have lower rewards than costs for unsafe behaviours. Do you understand?

Again, you will get 50 cents for stopping before a train comes.
You will be charged $1.00 for stopping if there is no train, and charged $2.00 if you fail to stop for a train.

The bonus money does not effect the $5.00 you get for showing up and participating. It is above and beyond the $5.00.

The computer behind you will keep track of your bonuses and I will tell you at the end of each block as to how much money you have earned.

Today you will interact with 4 blocks, the first two are for practice. Each block takes approximately 10 minutes to complete. After each block I will ask you to complete a short questionnaire. As the first two blocks are for practice these will not count in the bonus money. Do you have any questions?

Should we start with the first practice block? Remember you need to press the accelerator to start.

START FIRST PRACTICE BLOCK AND EXPLAIN EVENTS IF DEEMED NECESSARY.
CONTINUE WITH REMAINING BLOCKS, WHILE ADMINISTERING APPROPRIATE QUESTIONNAIRES.

TELL THEM ABOUT THE MONEY EARNED ONLY AFTER THE QUESTIONNAIRE HAS BEEN COMPLETED.

AT END WHEN ALL BLOCKS COMPLETED.

IF THE PARTICIPANT APPEARS DISTRESSED THEN REFER THEM TO :

1) Dr. Jeff Caird
2) University of Calgary Counseling Services

Appendix H
Critical event order in the four blocks: Study 2

83% reliability

	Event 1	Event 2	Event 3	Event 4	Event 5	Event 6
Block 1	4 *(Valid)*	8 *(Valid)*	13 *(Valid)*	15 *(Valid)*	18 *(Valid)*	22 *(Valid)*
Block 2	2 *(Valid)*	6 *(Valid)*	11 *(Valid)*	14 (Valid)	17 *(Valid)*	20 *(Valid)*
Block 3	3 *(Valid)*	5 *(Valid)*	9 **(Invalid)**	12 *(Valid)*	16 *(Valid)*	21 *(Valid)*
Block 4	5 *(Valid)*	7 *(Valid)*	10 *(Valid)*	13 *(Valid)*	18 *(Valid)*	23 *(Valid)*

50% reliability

	Event 1	Event 2	Event 3	Event 4	Event 5	Event 6
Block 1	4 *(Valid)*	8 *(Valid)*	13 *(Valid)*	15 *(Valid)*	18 *(Valid)*	22 *(Valid)*
Block 2	2 *(Valid)*	6 *(Valid)*	11 *(Valid)*	14 (Valid)	17 *(Valid)*	20 *(Valid)*
Block 3	4 **(Invalid)**	7 *(Valid)*	10 **(Invalid)**	12 **(Invalid)**	15 *(Valid)*	19 *(Valid)*
Block 4	3 *(Valid)*	8 *(Valid)*	12 *(Valid)*	16 *(Valid)*	19 *(Valid)*	22 *(Valid)*

Appendix I
Railway crossings for four blocks: Study 2

Block 1 for both reliabilities:

#	Location	Season	Event
1	Lions	Winter	
2	Thirty-Two	Winter	No Crossing
3	Erlton	Summer	
4	Fdeep	Summer	Event (True)
5	Gfast1	Winter	No Crossing
6	Sunny	Summer	
7	GlenEast	Summer	
8	GlenWest	Winter	Event (True)
9	Ramsay	Summer	No Crossing
10	Four	Winter	
11	Deerfoot	Summer	No Crossing
12	Fdeep	Winter	
13	NineOn	Summer	Event (True)
14	Fhuba	Summer	
15	Erlton	Winter	Event (True)
16	TenSt	Winter	
17	IngleD	Winter	No Crossing
18	Lions	Summer	Event (True)
19	Fhuba	Winter	
20	NineOn	Winter	
21	Nine	Summer	No Crossing
22	GlenEast	Winter	Event (True)
23	Barlow	Summer	
24	IngleD	Summer	

Block 2 for both reliabilities:

#	Location	Season	Event
1	Lions	Winter	
2	IngleD	Summer	
3	GlenEast	Summer	Event (True)
4	Nine	Summer	No Crossing
5	Fhuba	Winter	Event (True)
6	Erlton	Summer	
7	Deerfoot	Summer	No Crossing
8	Four	Winter	
9	Fdeep	Summer	Event (True)
10	Tenst	Winter	
11	Gfast1	Winter	No Crossing
12	Lions	Summer	Event (True)
13	Thirtytwo	Winter	No Crossing
14	GlenEast	Winter	
15	Fhuba	Summer	
16	GlenWest	Winter	Event (True)
17	Four	Summer	
18	Ingled	Winter	No Crossing
19	Fdeep	Winter	
20	Ramsay	Summer	No Crossing
21	Erlton	Winter	Event (True)
22	NineOn	Winter	
23	GlenWest	Summer	
24	Sunny	Summer	

Block 3 for 83% reliability:

#	Season	Event	Cond. 1 (FA1)	Cond. 2 (FA2)	Cond. 3 (MS)
1	Summer		NineOn	NineOn	NineOn
2	Winter	Event (True)	Four	Four	Four
3	Winter		GlenWest	GlenWest	GlenWest
4	Summer	No Crossing	Deerfoot	Deerfoot	Deerfoot
5	Winter		Fdeep	Fdeep	Fdeep
6	Summer	Event (True)	Fhuba	Fhuba	Fhuba
7	Summer	No Crossing	Nine	Nine	Nine
8	Summer		IngleD	NineOn	IngleD
9	Winter	Event (False)	NineOn	BarlowO (sum.)	NineOn
10	Summer	No Crossing	Ramsay	IngleD	Ramsay
11	Summer		Barlow	Ramsay	Barlow
12	Winter	Event (True)	Lions	Lions	Lions
13	Winter		GlenEast	GlenEast	GlenEast
14	Summer		Sunny	Sunny	Sunny
15	Winter		Erlton	Erlton	Erlton
16	Winter	No Crossing	IngleD	IngleD	IngleD
17	Summer	Event (True)	GlenWest	GlenWest	GlenWest
18	Winter		Tenst	Tenst	Tenst
19	Winter	No Crossing	Gfast1	Gfast1	Gfast1
20	Summer	Event (True)	Erlton	Erlton	Erlton
21	Summer		GlenEast	GlenEast	GlenEast
22	Winter		Fhuba	Fhuba	Fhuba
23	Winter	No Crossing	Thirtytwo	Thirtytwo	Thirtytwo
24	Summer		Lions	Lions	Lions

Block 4 for 83% reliability:

#	Location	Season	Event
1	Deerfoot	Summer	No Crossing
2	Erlton	Winter	
3	GlenWest	Summer	
4	Thirtytwo	Winter	No Crossing
5	NineOn	Summer	Event (True)
6	Lions	Summer	
7	Fdeep	Winter	Event (True)
8	Ingled	Winter	No Crossing
9	Tenst	Winter	
10	GlenEast	Winter	Event (True)
11	Barlow	Summer	
12	Fhuba	Summer	
13	Lions	Winter	Event (True)
14	GlenEast	Summer	
15	Ramsay	Summer	No Crossing
16	GlenWest	Winter	
17	Fdeep	Summer	
18	Four	Summer	Event (True)
19	Lions	Winter	
20	Gfast1	Winter	No Crossing
21	NineOn	Winter	
22	Nine	Summer	No Crossing
23	Erlton	Summer	Event (True)
24	Fhuba	Winter	

Block 3 for 50% reliability:

#	Season	Event	Cond. 1 (FA1)	Cond. 2 (FA2)	Cond. 3 (MS)
1	Summer		Fdeep	Fdeep	Fdeep
2	Summer		Lions	Lions	Lions
3	Winter		NineOn	NineOn	NineOn
4	Winter	Event (False)	IngleD	Gfive	IngleD
5	Summer		Sunny	Sunny	Sunny
6	Winter		Fhuba	Fhuba	Fhuba
7	Summer	Event (True)	Four	Four	Four
8	Summer	No Crossing	Ramsay	Ramsay	Ramsay
9	Winter		Lions	Lions	Lions
10	Summer	Event (False)	Fhuba	Sunnyside	Fhuba
11	Summer		NineOn	NineOn	NineOn
12	Summer	Event (False)	Erlton	Nine	Erlton
13	Winter		Four	Four	Four
14	Summer	No Crossing	Nine	Nine	Nine
15	Winter	Event (True)	GlenEast	GlenEast	GlenEast
16	Winter	No Crossing	ThirtyTwo	ThirtyTwo	ThirtyTwo
17	Winter		GlenWest	GlenWest	GlenWest
18	Summer		Barlow	Barlow	Barlow
19	Winter	Event (True)	Fdeep	Fdeep	Fdeep
20	Summer		GlenEast	GlenEast	GlenEast
21	Winter	No Crossing	Gfast1	Gfast1	Gfast1
22	Winter		Tenst	Tenst	Tenst
23	Winter		Erlton	Erlton	Erlton
24	Summer	No Crossing	Deerfoot	Deerfoot	Deerfoot

Block 4 for 50% reliability:

#	Location	Season	Event
1	Erlton	Summer	
2	Gfast1	Winter	No Crossing
3	Lions	Winter	
4	NineOn	Winter	Event (True)
5	Nine	Summer	No Crossing
6	Erlton	Winter	
7	GlenEast	Summer	Event (True)
8	Ramsay	Summer	No Crossing
9	GlenEast	Winter	
10	Fdeep	Summer	Event (True)
11	Fhuba	Summer	
12	IngleD	Winter	No Crossing
13	NineOn	Summer	
14	Barlow	Summer	
15	Fdeep	Winter	
16	Lions	Summer	
17	Thirtytwo	Winter	No Crossing
18	Four	Winter	Event (True)
19	GlenWest	Summer	Event (True)
20	Tenst	Winter	
21	Four	Summer	
22	Fhuba	Winter	Event (True)
23	Deerfoot	Summer	No Crossing
24	GlenWest	Winter	

Appendix J
Post-experiment debriefing protocols

Study 1

Thank you very much for your participation. The purpose of this Study was to select locations where HUD warnings about approaching trains should appear. Your assistance has helped us.

Pay the participant and obtain signature for remuneration.

Remind the participant that the below information is on their consent form and they are welcome to contact us for results.

Study 2

Thank you very much for your participation. The purpose of this Study was to test how HUDs effect braking to crossings and how false alarms effect trust in the system. Your assistance has helped us.

Pay the participant and obtain signature for remuneration.

Remind the participant that the below information is on their consent form and they are welcome to contact us for results.

Advance in-vehicle warnings for railway crossings

If you have any concerns or questions, or are interested in the results, please contact Jasdeep Chugh at 220-5910 or Dr. Jeff Caird at 220-5571. If you have further questions regarding your participation in this project you may also contact Dr. T. B. Rogers at 220-6378.

Appendix K

Participant responses to post-experiment questionnaire: Study 2

(1) Do you think the auditory and visual warnings expressed an appropriate level of urgency?

- The auditory plus visual aided in making it more salient a warning.
- I found that I was regularly stopping well in advance of the train.
- Bells close to real life - heard them 1^{st} then saw train. Sign came on at an appropriate time frame to allow you to stop in lots of time.
- Because it showed that a train was coming and to slow down right away.
- Yes, nice, loud and repetitive.
- Yes, they were accurate predictions of the expected situation, subtle enough of to be noticed in appropriate time for action.
- Yes, the auditory warning especially expresses urgency, more so than the visual as it is quite a loud, shocking sound.
- Yes, made more aware of the oncoming train.
- Yes, because the auditory signal was hard to ignore and furthermore it made you more alert as to the fact that you were approaching a railway crossing.
- ~~Yes, sound of train warning same as real life, makes you think really coming.~~
- Yes, tone was appropriate. Loud and quick.
- Yes, it wasn't alarming or shocking, but subtle and to the point.
- Yes, wasn't too distracting but did catch your attention to slow down.
- Yes, it gave you enough time to stop the car.
- No, it only makes the sound of a train approaching, which we would hear anyway. It should be more like an alarm. The visual sign is appropriate.
- Yes, provided adequate time to stop.
- Yes, the color yellow stands out against all other background colors and the noise is very cacophonous - but not as intrusive or urgent as ambulance etc.
- Yes, visible sign and loud bells.
- Yes, they provided early warning, sometimes too early. It provided the proper sense of urgency.
- Yes, auditory warnings were clear and the tone was urgent. Visual warnings were also clear.
- Yes, the bells were high pitched. Appropriate to get people's attention. Train can kill. I want urgent warnings.
- No, signal could have appeared for a longer amount of time.
- Yes, It gave you enough time to react.
- Yes, because the sign coincides to the highway warning of train tracks and the auditory warning reinforces the visual.
- Yes, nice chimes.

(2) Was the choice of the visual sign appropriate?

- It was the same sign as on the road traffic signs.
- It seemed to be a standard North American icon for railway crossing.
- Did not really look at sign but yellow worked and noted it was some RR related sign.
- It clearly indicates a railway sign (similar to road sign) and for people that can't see well it has auditory effects.
- Yes, it was bright.
- Yes, standard recognized symbol with easily discernible meaning.
- Yes, the yellow especially catches the eye, which is important. The clear and concise train track symbol allows you to quickly realize what the warning is.
- Yes, corresponds with the road sign.
- Yes, because it shows a simple yet highly practical insignia to make the driver aware of the danger. Because it is simple it is quickly associated with the message it is designed to deliver.
- Yes, could have been a little large.
- Yes, bright, appropriate sign.
- Yes, illustrates train, see train sign everyday.
- Yes, because it's the same as a street sign.
- Yes, easily interpretable as a railway crossing.
- Yes, it alerts you.
- Yes, it gave you proper notice that you were approaching a railway stop and a train was coming.
- Yes, yellow is a bright color. The sign was self explanatory.
- Yes, clear, symbol easily recognizable.
- Yes, yellow train tracks are already associated with train in my previous driving experience.
- Yes, tracks.
- Yes, it signified a train approaching.
- Yes, bold, yellow, shaped as typical warning sign.
- Yes, it's yellow. Yellow signs warn you of something ahead.
- Yes, it is a familiar symbol for trains.
- No, I think it could have been larger and sometimes one's mind is else where when they are driving (I'm speaking for myself).

(3) Was the choice of the auditory warning appropriate?

- It could have been bit choppy - it seemed it had a glitch.
- It is the same warning as the standard crossing guards.
- Bells of train very appropriate. Related to what was approaching.
- It symbolized approaching train.
- Sounded like a train.
- Distinctive, unlikely to be otherwise part of the environment.

- Yes, reference to a railway track stopping is appropriate, whereas some other sound (unrelated) would make no sense.
- Yes, it's abrupt and definitely catches your attention.
- Yes, it's about a train.
- Yes, the bell signal is a familiar sound that is also quickly associated with the message that is being delivered.
- Yes, maybe a little louder.
- Yes, sound of train coming.
- Yes, because it catches one's attention.
- Yes, it is the same as the actual noise.
- No, a little bit irritating to listen to although it did get the message across.
- No, they got annoying after a while.
- No, should be a voice instead saying, "train".
- Yes, reflected train's sounds - meaning was evident.
- Yes, ringing bells are already associated with trains.
- Yes, it is what people would associate of a train.
- Yes, it sounded very out of the ordinary and induced fear or sense of trouble. Hard to ignore.
- Yes, bells sound like bells that you hear at a railway crossing.
- Yes, the bells were high pitched. Appropriate to get people's attention.
- Yes, it sounds like the railway crossing.
- Yes, good combination to visual.

(4) Did the in-vehicle warnings come on at the right time?

- No, maybe a little too soon - didn't really feel the need to brake right away.
- Yes, it allowed for a smooth slow deceleration. Safer for winter operations.
- Yes, some times too early in experiment but feel would be OK in car as in car can gauge your threshold for braking. I would note sign then gauge distance before braking.
- Yes, because it gave you optimal time to understand what's going on and act towards it.
- No, sometimes they came too soon.
- No, because I ended up stopping before the crossing. It left me wondering if I could have made it.
- No, of course it depends on speed/road conditions etc. but they seemed to come on a bit soon. However, it is good to be over cautious.
- No, I would say it came too early.
- Yes, most of the times I felt it came on at the right time, thereby giving you enough time to brake gently.
- Yes, I felt I had enough time to stop.
- Yes, for me it was perfect.
- Yes, it gave me appropriate time to react, slow and stop.
- No, too soon. Always felt like I had too much time to stop.

- No, sometimes they came too late.
- No, too soon.
- Yes, adequate stopping time.
- No, they came too soon. I had to brake once - move forward and then brake again.
- Yes, plenty of time to stop.
- No, it came slightly too soon.
- Yes, enough time to brake safely.
- Yes, it gave me time to look for railway crossing and judge my stop appropriately.
- Yes, sufficient time to slow down.
- No, I think they could have been a little later due to the discrepancy between the warning ad the actual coming train.
- No, too soon. I would brake and then no train would be visible.

(5) Can you suggest design improvements to the in-vehicle warning system?

- Have a countdown system saying train in 500 m, 400 m, 300 m, etc.
- It could also be incorporated with a separate icon set for emergency vehicles. Direction that the train is approaching would also be an asset.
- Why on the left hand side of drivers view? Perhaps right hand side better or allow driver to choose - sometimes vision on one side better than other.
- Maybe change color to red to symbolize warning.
- Sometimes it went off, but there was no train.
- Maybe something using zone of safe travel ideas - more continuous rather than discrete levels.
- The visual aspect could be made stronger. The visual should be placed where the driver feels comfort (be given choice). Visual could blink to emphasize importance.
- You could place the visual sign in the middle of the screen rather than the bottom part. Maybe some drivers won't pay attention if it's at the bottom part.
- Perhaps is the signal was designed to not just come on only at times when a train was coming, rather it would come on at all active railway crossings. In today's urban areas there are not many active railway crossings, but lots of abandoned ones and I have seen many people not even slow down at any railway crossings. Therefore, if a warning system could discern between active and non-active crossings, that would prove to be very helpful.
- Larger visual warning, louder audio.
- Make the warning sign stay on until train leaves.
- HUD should come up depending on one's speed. Faster one is going, the earlier the warning.
- Warning system was good but how well would auditory warning work in noisy driving environment (e.g., music, engine, etc.)? The experiment kind of forces people to trust system because there are no other cues in the environment (e.g., flashing lights, actual train in periphery).
- The timing of the warning (which would also be different based on how slippery the roads are).

- Sign on driver's side?
- Maybe change color to red which may be easier to see when looking into the sun or purple which a happier color.
- Perhaps a pre-warning like the visual flasher, then when stopping time is very crucial, turn on the auditory warning.
- Back-up warning if there is no braking after a certain time period, like a safety feature.
- Bigger sign perhaps. Have auditory warning override the car's stereo system.
- When I saw the signal, I braked. Then the clip stopped. When I accelerated again, the train clip started. Then I braked again. It seemed like I didn't brake in time, but I did.
- Make it more reliable.
- Perhaps larger. Red in color.
- Just put the warning thing closer to when the real train comes.

(9) What would you do if the failures occurred on a regular basis?

False Alarm 1
- Not trust it but still follow it because better safe than sorry.
- I would turn the system off.
- Probably it would become habit to not trust the warning.
- Would not trust the system.
- Turn it off.
- Possibly lose faith in system and start to ignore it.
- Depend on other external cues.
- I'd ignore / dismantle the warning system.
- Stop using it.
- Check the system.

False Alarm 2
- Ignore, turn off the warning system.
- It would still give a warning that you are approaching a railway crossing - speed should be reduced and caution exercised. I would use the warning to slow and look for trains.
- Stop listening to the warnings and use my own judgment.
- Attempt to disable the system.
- Get rid of the warning system, may start to ignore it.
- Not trust it.
- I wouldn't trust it.
- Rely more on peripheral vision. Still slow down somewhat but keep going.
- Not believe in the system, might not brake nearly as quickly.
- This would be annoying.

Missed Signal
- I would watch lights vs. the sign if worked only half time would look for second cue to back sign up.

- Disconnect the warning system.
- I would try to have it checked out by a mechanic as soon as possible.
- Absolutely no trust.
- I would not trust the system.
- Trust my own judgments and look for trains.
- I would always slow down at train crossings.
- Always slow down a bit.
- Lose complete trust in the system.
- The signal system is a failure. I wouldn't trust it anymore.

(10) on an infrequent basis?

False Alarm 1
- Not trust it but still follow it but brake less harder.
- I would use it but use extra precautions.
- Failure on only 10-20% of the time wouldn't matter as much. I would trust the system more.
- Would not use it.
- Use it with caution.
- Would pay attention to it but would pay closer attention to salient cues in environment.
- Depend on other external cues.
- I'd be frustrated, but I'd keep it.
- Pay more attention to the warning by looking myself.
- Trust it, maybe. Try to get it fixed.

False Alarm 2
- again, turn it off.
- stop to ensure there was no oncoming train.
- I would pay attention and proceed cautiously.
- Use it as a backup to normal visual means of assessing the situation.
- Nothing, keep braking regardless.
- Be careful to check before reacting to stop.
- I wouldn't trust it.
- Rely more on peripheral vision. Still slow down somewhat but keep going. But not as much as if the failures occurred half of the time.
- Still have good faith in the warning system. O.K. (acceptable).

Missed Signal
- Same as above (watch for lights) but would rely more heavily on outside cues.
- Disconnect the warning system.
- Nothing, it would not bother me or at least I would try not to let it bother me.
- Absolutely no trust.
- I would not trust the system.

- Trust my own judgments and look for trains.
- I would always slow down at train crossings.
- Always slow down a bit.
- Lose complete trust because this is a life threatening situation so such warning is expected to be 100% reliable.
- That's still a high failure rate. I trust that there's a train coming when the signal appears, but if it doesn't appear, train might or might not be there.