Alarm effectiveness in driver-centred collision-warning systems

R. Parasuraman, P. A. Hancock† and O. Olofinboba†
Cognitive Science Laboratory, Catholic University of America,
Washington, DC 20064, USA
†Human Factors Research Laboratory, University of Minnesota,
Minneapolis, MN 55455, USA

Keywords: Alarms; Collision-warning systems; Driver reaction; Intelligent Travel Systems (ITS); False alarms; Posterior probability; Human-centred design.

The potential use of systems that seek to communicate a warning of impending collision directly to the driver is examined. Technological advances in collision-warning systems include reliable, low-cost radars, sensors with low noise levels, and the development of accurate detection algorithms for particular crash types, e.g. rear-end collisions. However, fundamental practical constraints make perfect sensor detection difficult to achieve. Imperfect detection conflates the false alarm rate and experience with other technologies confirms driver aversion to false warnings. Although sensitive alarm systems with high detection rates and low false alarm rates have been developed, the posterior probability of a collision given an alarm can be quite low because of the low base rate of collision events. As a result, only a small proportion of alarms will represent true collision scenarios. These and other factors can conspire to reduce alarm effectiveness in collision-warning systems. The problem is illustrated analytically and potential solutions are advanced.

1. Introduction

Road traffic accidents continue to be a major cause of loss of life and disabling injury throughout the world. In the USA, the estimated cost of motor-vehicle accidents exceeds $167 billion a year. In recent years a number of new technologies, collectively known as intelligent transportation systems (ITS), have been proposed to reduce accidents and to improve safety. Collision-warning and collision-avoidance systems are among the technologies that have been developed. Many have already entered the marketplace, e.g. the VORAD system used in buses and trucks in the USA, while other more advanced options are being planned under the rubric of ITS. In previous work (Hancock and Parasuraman 1992, Hancock et al. 1993, Hancock et al. 1996) the authors have advocated a driver-centred approach to the design of these systems, cognizant of the problems that have arisen in other automated systems that have excluded human operators from critical control loops (Parasuraman and Mouloua 1996).

The driver-centred approach requires that relevant information is brought to the attention of the individual in order that critical decisions directly involve the human operator. However, is this always possible? Comparable collision-avoidance systems such as the Traffic Alert and Collision Avoidance System II (TCAS II) have been introduced in commercial aviation and in maritime operations. In these transportation domains, such operator-based decision making is feasible given the time constraints involved and the extensive training that the operators of these systems receive. However, ground vehicles
do not permit the luxury of extended processing time and warnings must be passed immediately to untrained individuals to even countenance effective avoidance actions.

There are numerous issues that must be addressed before driver-centred warning systems that are effective and usable can be fielded. In this paper one of the major issues is examined, namely the nature and quality of the warning information provided to the driver. In particular the problem of false alarms is discussed and an analysis is presented of the problem based on the posterior probability of collision likelihood given an alarm.

2. **Collision detection algorithms**

In research and development work on collision-warning systems, first-pass engineering approaches have naturally concentrated on what appears to be one crash configuration directly amenable to reduction: rear-end collisions. These constitute about 23% of all crashes that are reported to the police and account for about 5% of all collision fatalities (Knippling et al. 1993a). A recent analysis of a controlled sample of such crashes found that the principal causal factor was ‘driver inattention’ (see also Rumar 1990); the next most important factor was ‘following too closely’ (Knippling et al. 1993b). Even given that other factors are likely to be involved, these analyses suggest that the rear-end crash rate could be reduced if an inattentive driver could be alerted in time. Alternatively, some crashes might be avoided if the driver could be cautioned not to follow so closely. A well-designed collision-warning system could potentially achieve both outcomes, but only if the driver could use the warning information in a timely and effective manner. As research on automated aids as shown, if the overhead involved in processing the information is high, then operators may choose (correctly) to ignore the aid (Kirlik 1993). A number of other crucial design issues have also to be considered before the anticipated benefits of collision-warning information can be realized. To begin with, for a collision-warning system to indicate potential collisions reliably and consistently, an appropriate detection algorithm must be used. Several possible algorithms have been proposed.

The simplest possibility is a time-to-headway criterion. This algorithm, which is applicable to rear-end collisions, refers to the time it would take a following vehicle to reach the lead vehicle or object, based only on the following driver’s speed. This is primarily applicable to situations where the lead vehicle is stationary (e.g. stopped at a traffic light, or stalled) and also in conditions of low visibility owing to poor weather. Although rear-end crashes into a stationary vehicle are more common than crashes where the lead vehicle is moving (Knippling et al. 1993a) a detection algorithm based only on a time-to-headway criterion would not be useful as a general criterion. More complex algorithms have been proposed that combine time-to-collision (TTC) measures based on headway and relative closing speed with other factors such as visibility and traffic density (Hogema and van der Horst 1994).

Farber and Paley (1993) carried out an interesting simulation analysis of rear-end collisions and examined the potential benefits that might accrue from a warning system using a detection criterion based on the relative closing rate of the lead and following vehicles. They applied a quasi Monte Carlo procedure to speed, closing rate, and headway data obtained from loop detectors installed on highway I-40 in Albuquerque, New Mexico, USA. Data for a total of 35,689 vehicle pairs were analysed. The detection criterion used was

\[ D = (FCS - LCS)^2 / 2a + (FCS - LCS)RT \]

where \( D \) = warning distance, \( FCS \) = following vehicle speed, \( LCS \) = lead vehicle speed,
$a =$ following vehicle deceleration, and $RT =$ following driver response time (to the onset of the lead car braking and to the collision warning signal).

Random values of $a$ and $RT$ were used in each iteration of the simulation to calculate the kinetics of the two vehicles braking to a stop. In this model, the warning signal is sounded if the actual distance between vehicles is less than the calculated warning distance $D$. Without collision warning, Farber and Paley (1993) estimated a crash rate of 173 for every million lead vehicle stops. With the warning system, they estimated that the crash rate could be reduced by at least 50%. Knipling et al. (1993b) carried out a similar Monte Carlo analysis using slightly different detection algorithms and estimated that between 37% and 74% of rear-end crashes could be theoretically prevented with a headway detection system.

The analyses carried out by Farber and Paley (1993) and by Knipling et al. (1993b) are important first steps in establishing figures of merit against which the effectiveness of different collision-warning systems can be compared. Clearly these analyses represent the theoretical upper-bound for effectiveness that practical systems should strive to reach. Whether a given warning system can achieve such levels of effectiveness requires a consideration of additional issues. As Farber and Paley (1993) themselves noted, such models raise but cannot answer certain critical human factors questions, such as the driver’s response to rare events like collision warnings, the impact of false alarms on driver behaviour, and driver acceptance of collision-warning systems. Although each of these issues is important, the authors focus in this paper on the false alarm problem, which contributes significantly to the overall effectiveness of warning systems and therefore to the related issues of driver response to and acceptance of these systems.

3. False alarms: setting decision thresholds for automated warning systems

Unfortunately, many proposed collision-warning systems are vulnerable to false alarms. The false alarm question is not trivial. As noted earlier, the comparable warning system in aviation, TCAS, suffered considerably in its early development from problems of excessive false alarm rates that resulted in mistrust and lack of pilot usage (Wiener 1988). Poorly designed warning and alerting systems in the cockpit promote not only disuse but also what Satchell (1993) engagingly referred to as ‘creative disableness’ by pilots. Solving the false and nuisance alarm problem will be similarly crucial for future ITS deployment (Horowitz and Dingus 1992, Knipling et al. 1993b).

The selection of a detection criterion must balance the need for early detection with the avoidance of false alarms. Signal detection theory (SDT) provides a basis for determining the appropriate decision threshold or bias that balances these two needs (Swets and Pickett 1982). If $S$ represents the environmental event or signal associated with a collision, $N$ a non-collision or noise event, and $R$ a positive alarm response of the warning system, then $P(R|S)$ and $P(R|N)$ are the probabilities of correct detections (hits) and false detections (false alarms), respectively. The accuracy of the detection system ($d’$) and the response bias ($\beta$) can be derived from these probabilities using standard SDT computing formulae; alternatives to $d’$ and $\beta$ are also available (MacMillan and Creelman 1990). These measures can then be used to evaluate the effectiveness of different detection algorithms, for example such as that proposed by Farber and Paley (1993). Furthermore, detection of collisions represents an instance of a joint human-computer monitoring system—both the human and the machine detector directly observe the environmental stimuli indicating possible collision, and the human makes a joint decision combining this input with the output of the warning system. SDT has been applied to the issue of determining the joint

The initial consideration for setting the decision threshold (β) of an automated warning system is the cost of a miss versus that of a false alarm. Missed signals (collisions) have a phenomenally high cost, yet their potential frequency is undoubtedly very low. Indeed some drivers may drive for decades without taking the sort of evasive action mandated by the proposed systems. However, if a system is designed to minimize misses at all costs, then the problem of frequent false alarms is immediately encountered. A low false alarm rate, and arguably a zero false alarm rate, would appear to be critical for acceptance of collision-warning systems by drivers. Accordingly, setting a strict or conservative decision threshold to obtain a low false alarm rate would appear to be good design practice. Yet, as is shown later, this may be insufficient by itself; the posterior probability of collision likelihood given an alarm must also be considered.

Should the decision threshold be set as strictly as possible? Perhaps not, because the failure to supply sufficiently advanced warning could be equally problematic. Farber and Paley (1993) suggested that too low a false alarm rate may also be undesirable, given that police-reportable rear-end collisions are very rare events (perhaps occurring once or twice in the lifetime of a driver). If the system never emits a false alarm, then the first time the warning sounds would be just before a crash. The driver may not respond quickly to such an infrequent event. Farber and Paley (1993) speculated that an ideal detection algorithm might be one that gives an alarm in collision-possible conditions, even though the driver would probably avoid a crash. Although technically a false alarm, this type of information might be construed as an aid in allowing improved response to an alarm in a collision-likely situation. Thus all false alarms need not necessarily be harmful. This idea is similar to that of graded warnings, and to the concept of ‘likelihood’ displays as espoused by Sorkin et al. (1988). Either method of signalling possible collisions might be particularly useful in training novice drivers to recognize appropriate following distances when driving without a warning system. For the experienced driver, however, the only acceptable system would be one with a true false alarm rate that is indistinguishable from zero.

4. Alarms and collisions: the importance of the posterior probability of a true alarm

For any collision-warning system, d' will be determined by such factors as sensor noise and the effectiveness of the detection algorithm, whereas β can be freely set by the system designer to achieve different hit and false alarm rates for a system with a given value of d'. The technology exists for system engineers to design sensitive warning systems with high values of d'; and these systems will generally be set with a decision threshold that minimizes the chance of a missed warning while keeping the false alarm rate below some low value. The importance of keeping the false alarm rate as close to zero as possible has already been indicated previously. However, despite the best intentions of designers, the availability of the most advanced sensor technology, and the development of very sensitive detection algorithms, one fact may conspire to limit the effectiveness of alarms: the low a priori probability or base rate of potential collision scenarios. For the driver, that a given warning system may have a very high hit rate P(R|S) and a very low false alarm rate P(R|N) is less important than the inverse of these conditional probabilities, i.e. the posterior probabilities P(S|R) and P(N|R). In particular, the posterior probability P(S|R) is of vital importance to the human monitor of any automated alarm system. This represents the probability that a given alarm response represents a true condition (e.g. a
potential collision). Bayes’ theorem can be used to derive \( P(S|R) \):

\[
P(S|R) = \frac{P(R|S)}{[P(R|S) + P(R|N)(1 - p)]/p}
\]

where \( P \) is the \textit{a priori} probability of the signal \( S \).

If the base rate \( p \) is low, as it often is for many real events, then the posterior odds of a true alarm can be quite low even for very sensitive warning systems with very high hit rates and low false alarm rates. Figure 1 plots posterior probability values as a function of \textit{a priori} probability for a sensitive detection system of given accuracy \( d' = 4.65 \). The family of curves represents the different posterior probabilities that can be achieved with different decision criteria \( \beta \). Figure 2 shows a similar set of curves for an even more sensitive detection system having \( d' = 6.18 \). As these figures indicate, the posterior probability \( P(R|S) \) approaches 1.0 only for relatively high values of the \textit{a priori} probability \( p \). Conversely, for low values of \( p \), the posterior probabilities are quite low even for a very sensitive system.

These figures can perhaps be better appreciated by computing the odds for specific cases. Assume a sensitive collision-warning system with a \( d' \) of 4.65 and assume that the decision threshold \( \beta \) is set to 1. This gives a near-perfect hit rate for the system of 0.99 and a false alarm rate of only 0.01. Despite these impressive detection statistics, the human operator could find that the posterior odds of a true alarm with such a system can be quite low. Table 1 shows the values of the posterior probability \( P(S|R) \) and the associated odds of a true alarm for several values of \( p \) ranging from 0.0001 to 0.1. As table 1 shows, even

![Figure 1](image.png)

**Figure 1.** Posterior probability \( P(S|R) \) of a signal (collision event \( S \)) given an alarm response \( R \) for a collision-warning system with a given sensitivity \( d' \) of 4.65, plotted as a function of the \textit{a priori} probability \( p \) (base rate) of \( S \). The parameter that generates the family of functions is the decision criterion \( \beta \); functions 1–4 are associated with progressively lower (more liberal) values of \( \beta \). The corresponding hit \( P(R|S) \) and false alarm \( P(R|N) \) rates are also shown.
with such a sensitive detection system, when the a priori probability is low, say 0.001, only 1 in 11 alarms that the system emits represents a true alarm. The situation is a little better if the decision threshold is set a little stricter, sacrificing hit rate for a lower false alarm rate. Now 3 out of every 7 alarms represent a true alarm for a \( p \) of 0.001. If the designer is tempted to use a more liberal decision threshold in order to achieve an almost perfect detection rate of 0.999, then when \( p \) is 0.0001, the posterior odds of a true alarm are whittled down to as low as 1 in about 600! Little wonder then, that many human operators tend to ignore or to turn off alarms—they have cried wolf once too often (Sorkin 1988); or, as in the actual case of a prison escape in which the escapee deliberately set off a motion detector alarm, they are slow to respond to the alarm (Casey 1993).

As figures 1 and 2 and table 1 indicate, consistently true alarm response occurs only when the a priori probability of a signal is relatively high. There is no guarantee that this will be the case in many real systems. More generally, these results suggest that designers of collision-warning systems must take into account both the decision-threshold at which these systems are set and the a priori probabilities of the collision types, if these are known. Swets (1992) has discussed procedures for effective choosing of appropriate values of \( \beta \) for diagnostic systems. Decision thresholds for collision-warning systems must be set not just for high hit and low false alarm rates, but for relatively high values of posterior true alarm probabilities as well.

The importance of posterior probabilities can also be seen if the data from the Farber and Paley (1993) study are used. They estimated a base rate of 173 crashes for every million stops, or a \( p \) of 0.000173. Given this base rate, the detection system shown in table
Table 1. Values of the posterior probability \( P(S|R) \) of a signal event (S) given a alarm detection response (R), and the associated odds of a true alarm, for several values of the a priori probability \( p \) for a detection system of fixed sensitivity, \( d' = 4.65 \).

| \( \beta \) set to 1 | \( P(R|S) = 0.99 \), | Posterior probability \( P(S|R) \) | Odds of a true alarm |
|---------------------|------------------|-----------------------------|-------------------|
|                     | [\( P(R|N) = 0.01 \)] |                             |                   |
| 0.0001              | 0.0098           | 1 in 102                    |                   |
| 0.001               | 0.09             | 1 in 11                     |                   |
| 0.01                | 0.5              | 1 in 2                      |                   |
| 0.1                 | 0.92             | 10 in 11                    |                   |

| \( \beta \) set to 24 | \( P(R|S) = 0.95 \), | Posterior probability \( P(S|R) \) | Odds of a true alarm |
|-----------------------|------------------|-----------------------------|-------------------|
|                       | [\( P(R|N) = 0.0013 \)] |                             |                   |
| 0.0001               | 0.068            | 1 in 15                     |                   |
| 0.001                | 0.42             | 3 in 7                      |                   |
| 0.01                 | 0.88             | 8 in 9                      |                   |
| 0.1                  | 0.988            | 100 in 101                  |                   |

| \( \beta \) set to 0.029 | \( P(R|S) = 0.999 \), | Posterior probability \( P(S|R) \) | Odds of a true alarm |
|--------------------------|------------------|-----------------------------|-------------------|
|                          | [\( P(R|N) = 0.0594 \)] |                             |                   |
| 0.0001                   | 0.068            | 1 in 596                    |                   |
| 0.001                    | 0.017            | 1 in 60                     |                   |
| 0.01                     | 0.145            | 1 in 7                      |                   |
| 0.1                      | 0.651            | 2 in 3                      |                   |

1, for which \( d' = 4.65 \), would yield posterior odds of a true alarm of only 1 in 60. Even for an extraordinarily sensitive system, with a hit rate of 0.9999 and a false alarm rate of 0.0001 \( (d' = 7.4) \), the posterior odds of a true alarm would be only about 2 in 3. These numbers attest to the powerful influence of the base rate on the true alarm rate. Farber and Paley (1993) concluded that with the use of a collision-warning system the crash rate could be reduced, in theory, from 173 per million stops to less than half that amount. They were, however, aware that this represents a best-case scenario. The present analysis suggests that any system in practice (that includes a human element) would be hard pressed to achieve this level of reduction in crash rate.

5. Related considerations

Determining alarm effectiveness rests not only on assessment of posterior probabilities of collision likelihood but also on several other factors as well. In a detailed assessment of research needs for in-vehicle collision avoidance devices, Hanowski et al. (1994) identified a dozen major issues that needed further investigation. Among the areas that they discussed were the effectiveness of visual, auditory, tactile, proprioceptive, and multi-model warnings, the spatial location of warning information, temporal sequencing of warning, and driving changes as a result of warning availability. The scope of the ergonomic challenge should be clear, because as yet only limited information is available for these and related questions. In this paper the authors have necessarily restricted their analysis to a few collision types, principally two-vehicle collisions occurring for vehicles travelling in the same or opposite directions. Many other collision types are possible, and a comprehensive assessment of collision-avoidance systems will need to take into account collision typology (Massie et al. 1993).

Empirical studies aimed at evaluating the wide range of relevant factors are still ongoing, and only a few are mentioned here. A few studies have evaluated different detection criteria for collision-warning systems. For example, Janssen and Thomas (1994) compared different collision avoidance detection criteria on vehicle-following performance under simulated normal and low visibility conditions. They found beneficial effects for a system that used a 4 s TTC criterion; after being triggered the system increased the
usual counter-force on the accelerator pedal, thus providing a rapid kinesthetic feedback signal to the driver. Drivers with this system had fewer short following headways (< 1 s) and lower overall driving speed. Schumann et al. (1993) evaluated the potential to modulate vehicle control dynamics by modifying the driver’s control input. To accomplish this, different tactile feedback formats were contrasted with auditory warnings as methods to signal to drivers whether or not to remain in their lane when attempting passing manoeuvres. They found that drivers were more responsive to proprioceptive cues (such as steering wheel vibration or force feedback, e.g. resisting the driver’s control input to change lanes) as compared to auditory warnings.

Even a collision-warning system that uses a sensitive and reliable detection algorithm, maximizes the posterior probability of collision likelihood, and uses appropriate information presentation formats, may nevertheless be less effective than anticipated because drivers may drive differently with the system than without it. The presence and occurrence of warnings may themselves change driver behaviour. For example, a conservative warning system seeking to avoid all collisions may be triggered frequently but the driver may react to these warnings and drive more cautiously so that false alarms are minimized. In essence, this is the antithesis of risk homeostasis conception (Wilde 1982). What is likely is that collision-warning systems will influence driver behaviour in general, not just on those occasions when the warning is appropriate; these influences must be understood for warning systems to be effective.

There are a number of additional considerations on top of the question of quantitative assessment of warning system algorithms. Should developers focus specifically on collision-detection systems such as rear-end collision protection (Knippling et al. 1993b) or should they employ some form of general protective envelope approach (Hancock 1993)? The case-specific circumstance is, of course, representative of the need to start somewhere. In addition, can or should alarm systems be individualized such that they respond to likely accident conditions for the pertinent driver age group? Developing a collision-warning system implies first the implementation of some complex, multi-array detection system. In itself this is a considerable engineering challenge. However, having derived a veridical warning system, its customization for consumption by drivers of widely different abilities is uniquely a human factors question. It is this arena that promises enhanced safety, yet also represents the most complex portion of ITS development (Hancock et al. 1996).

6. Conclusions

We are fixed on the horns of a dilemma. If we choose to develop completely automated collision-avoidance systems, then we are constrained to design them for very specific circumstances like vehicle-following. The reason for this constraint is that as possible environmental configurations proliferate, they rapidly overwhelm the computational capabilities of the system involved. For example, it is difficult to convey to automated systems which of two collision alternatives might be preferred. Given the choice of colliding with a wall or standing crops, the human is able to make what is considered to be the obvious decision. However, this is not obvious to a machine detection system in that it needs complex resident knowledge about the fundamental nature of bricks and corn. Does this mean that we should not pursue fully automatic collision-avoidance systems? Certainly not. However, intrinsic problems remain such as the spectre of responsibility in a chain of rear-end collisions involving such equipped vehicles. More importantly, it is critical to convey the limitations of such systems to drivers. Already in ITS route-guidance systems it has been observed that drivers ‘over-attribute’ intelligence to apparently
sophisticated systems. What automated systems can and cannot do compared with what their human operators think they can or cannot do is a major concern (Parasuraman and Mouloua 1996).

If automated systems are limited, should we move directly to hybrid approaches in which drivers are either warned of automated action or are left to make critical decisions? This approach is more compatible with the ‘envelope of protection’ strategy advocated by Hancock (1993). However, as we have shown in this paper, the low posterior likelihood of collision given an alarm is particularly problematic. Designers of collision-warning systems must take into account both the decision threshold and the base rate of the collision events in order for drivers to trust and use these systems. The problem of low posterior odds of a true alarm is compounded when the provision of different alarms for different potential collisions confounds the overall false alarm rate. Currently proposals have been put forward for as many as 20 to 30 different types of collision-detection and avoidance systems for different driving circumstances (e.g. side impact). Such developments pose a serious challenge to the need to provide effective and timely warning information to the driver. Safety is frequently touted as the great contribution of ITS in general and collision-warning systems in particular. Much remains to be done to deliver on that promise.

Acknowledgements
Preparation of this paper was supported by grants from NASA Langley Research Center (NAG-1-1296) to Raja Parasuraman and from the Minnesota Department of Transportation (Mn/DOT) and the Center for Transport Studies and the ITS Institute at the University of Minnesota to P. A. Hancock. Thanks to Jackie Duley for comments and help in preparation of the figures.

References
HANGOWSKI, R. J., BITTNER, A. C., ROWLEY, M. S., WILSON, J. C., RABY, M., BYRNE, E. A. and PARASURAMAN, R. 1994, An information clearinghouse conceptual design: a system for meeting crash avoidance research needs relevant to older drivers. Final report, Battelle Seattle Research Center, Seattle, WA.


