AGONISING OVER ASILS: Controllability and the In-Wheel Motor

M. Ellims*, H.E. Monkhouse*

**Protean Electric Ltd, UK, mike.ellims / helen.monkhouse@proteanelectric.com*

Keywords: ASIL, Controllability, Severity, Human Factors.

Abstract

In the automotive domain the standard ISO 26262 places significant emphasis on the assignment of Automotive Safety Integrity Levels (ASILs). In particular much of Part 3 of the standard is dedicated to the process that determines the three factors that contribute to the final assigned ASIL value: exposure, severity and controllability. In this paper we examine some of the issues that the authors have encountered during the development of an in-wheel electric motor and will argue that the perceived emphasis on ASIL ratings, in the context of developing a safe system, is misplaced and potentially counterproductive.

1 Introduction

As indicated above, significant proportion of ISO 26262 Part 3 is devoted to assigning Automotive Safety Integrity Levels (ASILs). The means by which these values are assigned is further expanded in Part 3 Annex B. However, rather than providing a process or methodology for determining these properties, the reader is presented with a simplified set of example tables. For instance, section B.2 which develops the concept of Maximum Abbreviated Injury Scale (MAIS) in some detail, leaves the process by which the severity rating should or perhaps could be derived as "*Accident statistics can be used to determine the distribution of injuries that can be expected to occur in different types of accidents*". The information provided relating to exposure is somewhat more helpful, but does not address issues such as how different factors could or should be combined. Similar observations can be applied to the examples provided for controllability ratings in section B.4; where a table of driving situation examples is given with assumptions about the corresponding control behaviours that would avoid harm. Somewhat less clear is how to build the evidence that forms the rationale for the controllability rating chosen.

In section 1 we review the "*item*" with which we are concerned, the hazards associated with it, and briefly review the lessons that can be learned from history. Section 2 summaries the factors that feed into the ASIL determination and section 3 discusses these in more detail; noting some weakness in the way these factors are defined and observes that controllability is the critical factor. Section 4 examines controllability in the context of the driver and examines in detail what can, and more importantly what cannot, be expected of the driver. These driver expectations are then discussed in the context of the development of the functional safety concept for the in-wheel motor application. In section 6 we conclude that, if the safety goals (high level safety requirements) are incorrect, then getting the ASIL wrong is irrelevant.

1.1 In-wheel motor technology

The concept of using in-wheel motors as a means of vehicle propulsion was first conceived in the late 19th century, with the first patent being registered in 1884. In 1897 Ferdinand Porsche raced a car that had electric wheel motors. Although Porsche's wheel motors were more efficient than the gasoline and diesel powered vehicles of the day, the much higher energy density offered by petroleum over batteries meant that higher power and range were more easily achieved using an internal combustion engine.

Today ever more stringent emissions targets are driving vehicle manufacturers to move to add hybrid and electric power trains to their existing vehicle fleets; a factor driving renewed interest in in-wheel motors.

Figure 1: Brabus 4WD vehicle (EV1) fitted with Protean Motors during fault injection testing (Photo D. Harty).

1.2 Hazards & risks

One would consider many of the hazards associated with the use of in-wheel motors to be the same as that of a conventional power train; the one complicating factor being the potential independent torque control that is possible across the vehicle axle. The potential to control the asymmetric torque across an axle has significant vehicle dynamics

benefits, but if it occurs in error then there exists the potential to produce an un-commanded yaw moment.

The driver will likely react to the un-commanded yaw in the same way as perhaps they would when the vehicle is hit by a wind gust, or when the vehicle 'pulls' to one side having hit water on the road. As with a wind gust or puddle, the driver may apply the brake, but their primary response will be to simply apply a steering input correction in order that the vehicle maintains the desired heading. Whether the driver's intervention leads to a successful outcome will ultimately depend on a number of factors, not least of which is the magnitude of the yaw moment; a subject to which we will return later.

1.3 What history tells us

When setting targets for innovative technologies, as well as considering what will be deemed acceptably safe, one should also be aware of customer perceptions. Automotive history is littered with examples where the public has lost confidence in a particular brand or technology as a result of bad publicity; however factually inaccurate that publicity later becomes.

The route that Antilock Brake Systems (ABS) took in order to become a widely accepted vehicle technology has been a long one, which required a process of careful social-technical planning by Bosch, in order to overcome the negative publicity that surrounded ABS as a technology and to make Bosch ABS the dominant antilock technology [11]. In order to achieve this, Bosch engineers found that not only did they have to 'sell' carefully the performance achieved by the product, but also they had to 'sell' the setting of the standards on which that performance was measured.

As with ABS, the automotive community has preconceived ideas about the capability of in-wheel motors, and like ABS careful consideration must be paid to the setting and measurement of performance targets. A fact that has been at the forefront of the authors' mind when classifying the risks associated with in-wheel motor hazards; particularly when attempting to assign a quantitative measure to controllability.

2 Automotive safety integrity levels

ISO 26262 defines Automotive Safety Integrity Level (ASIL) as the "necessary requirements of ISO 26262 and safety measures to apply for avoiding an unreasonable residual risk." [10] There are four levels of ASIL, with ASIL D being the most stringent level and ASIL A the least. The ASIL is determined by considering the impact of severity, probability of exposure and controllability, and is based on the functional behaviour of the system under evaluation.

2.1 Severity

The Standard requires that "*the severity of potential harm is assessed for each hazard that has been identified. With the potential for harm being assessed for each person potentially as risk; be that the driver or passengers of the vehicle causing the hazardous event, or other people potentially at risk such as cyclists, pedestrians or the occupants of other vehicles*".

2.2 Exposure

ISO 26262 defines probability of exposure broadly as "*The probability of exposure of each operational situation shall be estimated based on a defined rationale for each hazardous event*".

2.3 Controllability

The Standard requires that "*the controllability of each hazardous event, by the driver or other persons potentially at risk, shall be estimated based on a defined rationale for each hazardous event.*" It then goes on to note that "*the evaluation of the controllability is an estimate of the probability that the driver or other persons potentially at risk are able to gain sufficient control of the hazardous event, such that they are able to avoid the specific harm.*" On paper both statements appear relatively straightforward, but when one begins to consider what the quantitative measure of controllability for a given hazard might be, or how to generate statistically relevant test evidence, the task suddenly feels less straightforward.

3 The problem in context

There are two primary issues with the ASIL system. The first is that it dominates the text of the standard and is somewhat out of proportion to the potential effect on system safety – with the intention of the risk ratings (ASIL) being to remove "*unreasonable **residual** risk*" by requiring a process that has a high probability of detecting errors. However, the critical issue often ignored is that you have to know what an error looks like to recognise it. The second issue is that when assigning ASIL values establishing consistent parameter values can be problematic. This is the topic of the remainder of this section.

3.1 Severity

ISO 26262 includes this normative guidance regarding severity "*The severity of potential harm shall be estimated based on a defined rationale for each hazardous event*". This guidance and the associated notes are not particularly useful to a reader attempting to assign severity ratings. In hindsight what is missing from the standard is the *intent* of the original authors and a precise definition of "potential harm".

Part 3 Annex B gives some hints that severity ratings could or perhaps should be derived from accident data, but it is not explicitly stated. And when considering the examples presented in the tables the reader could be left with the view that it is *possibly* worst case outcomes that one should consider; for example assigning S3 (defined by ISO 26262 as "life-threatening injuries (survival uncertain), fatal injuries") for a pedestrian bicycle accident on a two lane road. If the

latter case is true then all failures that affect the dynamics of a moving vehicle should *perhaps* be assigned a severity rating of S3. However, if it is the former then a severity rating of S3 may not be possible. For example, Morris *et al*. [16] used a combination of STATS 19 and Co-operative Crash Injury Study (CCIS) data to estimate the total probability of accidents with different MAIS levels[1]. The combined MAIS 5 and MAIS 6 probabilities for all accidents are less than 10% and MAIS 3+ just exceeds this level (for all accidents). Likewise data from NHTSA [17] indicates that this also appears to be true in the USA.

3.2 Exposure

If all factors relating to exposure are taken into account then it may provide a potential mechanism to account for severity definition imprecision in the standard. In Annex B factors that are listed include the type of driving being undertaken and weather conditions.

Perhaps of more interest is what is not stated within the standard. For example, consider the following situation, driving at high speed on a motorway, in daylight, and with fine weather in a low traffic density. In the base case the exposure will necessarily be rated E4 "*highly probable*" (or possibly E3 "*medium probability*") for equipment associated with drive train control. However, what is not immediately apparent is how the exposure could or should be adjusted when considering the persons at risk. For the vehicle occupants it will remain at E4, and it is probably E4 for persons travelling in other vehicles, but for pedestrians and cyclists it is less clear.

In such situations it seems reasonable that the exposure should take into account the probability of encountering those types of road users. Outside of Cambridge[2] it is unlikely that cyclists would be encountered on our example motorway, so a low exposure rating of E0 or E1 is possible. What is not clear is how far this could or *should* be taken, and one has to be mindful of "salami slicing" down to a lower ASIL.

3.3 Controllability

At a first glance controllability appears intuitively obvious; with Part 3 Table B.4 describing various compensating driver actions. The option "*maintain intended driving path*" being particularly popular, as is "*brake to slow/stop vehicle*", which together account for 12 of the 14 suggested control actions, which is not a criticism (see below).

Obviously, the real world is much more complex. Minor accidents occur relatively frequently with the majority probably not being reported or recorded; at least not to the police. More serious accidents are much less frequent, and the actual occurrence rate is rather lower than perhaps public perception would suggest; with 80 deaths or seriously injuries (KSI) per billion miles in England in 2010 [3]. These totals are small but still significant.

Of relevance here is the fact that the vast majority of accidents involved vehicles that were functioning as intended. Vehicle defects are reported as a contributing factor in only 2% of accidents [17] with tyres and brakes accounting for almost all of the faults reported.

Given the above, what are the prospects for our driver? Is the likelihood of them controlling the hazardous event better or worse than might be expected from considering the accident data? This quandary is the subject of section 4.

3.4 The state of the art

As previously stated, three factors feed into the risk classification: severity, exposure and controllability, and for any given situation, the factors severity and exposure can be considered fixed.

Severity of an accident *that has occurred* and involves a vehicle is determined almost completely by the change in speed of that vehicle (delta V), the collision geometry and the conservation of momentum. Likewise exposure for the most part is not determined, but rather enumerated; at least for continuously active devices. This then leaves us with controllability as the main mechanism by which we can influence the risk associated with a device failure.

4 The driver and controllability

Dewar and Olson [4] characterise the road systems as comprising three major elements: the road, the vehicle and the driver, and state that "*the driver is the least understood*". This is especially true in emergency situations. For instance Leach [12] provides figures for human reaction to disasters where: between 10% and 20% will remain calm, the largest group of around 75% will be "stunned and bewildered" and exhibit impaired thinking, while the remaining 10% to 15% will exhibit "inappropriate" behaviour i.e. panic.

So, the discussion thus far would seem to indicate that our ability to rate controllability, other than category C3 (less than 90% of drivers can control or avoid harm), is bleak. However, as stated in section 3.3, it is probable that the large majority of what could become accidents are actually avoided. So what is going on?

4.1 Braking

Part 3 cites braking the vehicle as an example action taken by the driver to avoid harm. But indications are that people can be very poor at applying the brakes sufficiently when required: with widely variable reaction times ranging from 0.32 seconds to more than 4 seconds [25], inconsistent responses i.e. too little, too late or even the release of the

[1] MAIS; Maximum Abbreviated Injury Scale, 0 is no injury, 1 minor, 2 moderate, 3 serious, 4 sever, 5 critical and 6 untreatable or fatal.

[2] Cyclists are occasionally seen on the M11 motorway; usually coincident with " fresher week" at the local universities.

pedal [1], and a corresponding large number of rear end collisions.

However, we must evaluate this information in context. Firstly, we should remember that the accident data is a biased sample. That is, it does not record episodes where a braking manoeuvre was successfully executed. Secondly, we have to consider the driver's inputs and their reaction to those inputs *in detail*.

4.2 Humans and braking

The simplest model of the human braking response involves three steps: perception (seeing), cognition (thinking) and reaction (doing). During "normal" braking this model is correct and in this case the Brake Reaction Times (BRTs) are distributed towards the higher end of the distribution.

But what of the BRTs that fall towards the low end of the distribution? With reaction times lower than half a second something else must be occurring. From neuroscience we find that our initial responses are largely "instinctive", with responses that take less than 500ms being almost wholly "automated". After the initial "instinctive" response, the frontal cortex takes over and we start to consciously control our actions (or not).

The basic structure of what is often referred to as the startle response is laid out by Staal [21], who attempts to create a conceptual framework of human performance under stress, using the term "*evaluative reflex*" to describe the initial instinctive reaction.

Of importance for controllability is that the initial evaluation is very fast (100-250 ms) and the response is correspondingly fast (300-500ms). Moreover, the region of the brain most often associated with these rapid responses (the amygdala) is well connected to the major sensory visual, acoustic, kinaesthetic and vestibular inputs. For example, the driver can respond quickly to the appearance of a red light, with Green [7] stating that "*hitting a brake pedal in response to the flashed brake lights ahead is an example of a learnt response*". We learn to associate brake lights with a particular action we should take; what Crawford and Cacioppo term "*statistical learning*" [2]. The problem is that we do not in general learn emergency responses, aside from those we regularly encounter, thus when presented with emergencies our reaction may be suboptimal.

4.3 Implications

From the above we can conclude several things: firstly, given a known stimulus the initial responses will be largely reflexive. Thus if there is a specific, learnt and automated response to the stimulus, then the outcome will likely be favourable. Otherwise, if no such learnt response exists, then we will probably at best get a generic response which may not always be appropriate. For example, "freezing" in response to a sudden sound maybe effective in the African savannah, but is somewhat less effective when driving. However, this may partially explain why a large proportion of drivers do not brake during emergency situations [22]. It is not difficult to find personal experience of the problem. One of the authors when they first encountered an emergency stop signal comprising flashing amber lamps (at 4Hz) took at least a second to respond with the required cognition.

So after the initial response, what comes next? So far we have only discussed the first half second of the driver's response. After this the only thing we can say with certainty is that we can't be sure what will happen. Given the example above, we know that it could be a second or more before the driver takes effective action, resulting from purposeful cognition having taken place, and in many accident scenarios involving the components of the vehicle drive train a second or more will be too long.

4.4 Humans and in-wheel motors

The motivation that has led to the information summarised above was the urgent need to understand how a driver could reasonably be expected to react to an in-wheel motor failure.

Enumerating the effects of in-wheel motor failure on a vehicle is a relatively simple exercise. The immediate effect on the vehicle is an acceleration or deceleration, and as the torque disturbance is off centre a yaw is induced. The hazards *unintended acceleration* and *unintended deceleration* are relatively well understood within the power train section of the automotive community. The hazard *induced yaw* less so.

The obvious step to explore induced yaw was to examine situations where a yaw moment is externally induced. Work by Wierwille *et al*. [24], on driver reaction time to simulated wind gusts, strongly suggest that the response was "natural" and fast; though at the time the reasons *why* this might be so was not fully appreciated. Information on the effects of standing water was sought, but very little was found to be of use, aside from the fact that decelerations can be severe – for 1.5 inches of water the deceleration can be in the order of 1*g* [9].

The major advance in the authors' understanding came from the large study by Neukum *et al*. [18] who examined the response of steering superposition (offset) errors and determined a maximum tolerable level of yaw that appears to be valid across different types of vehicle. Neukum *et al*. [18] and Neukum [19] show that the driver's response to this kind of disturbance is fast (180 to 220ms) and hence natural and more importantly automatic. In addition this pair of studies supports the idea that the yaw rate limits are not dependant on the inducing mechanism; supporting evidence from Wierwille *et al*. [24]. A strong connection linking these two sets of results was found in a wind gust study by MacAdam *et al*. [14] which also produced a yaw rate limit close to 2.5 ^{o}/s at 150 kph for a vehicle unbalanced aerodynamically.

Initial results from vehicle testing (Figure 1) give no reason to expect that the driver's behaviour, in the presence of failures, does not match expectations from steering fault and wind gust data. Figure 2 shows a trace of a single event on the left front

wheel. The change in steering angle occurs around 0.2s after the event is triggered and the correction is complete at approximately 0.5s. The other interesting feature shown is the spike on the steering wheel torque sensor induced by the torque error. This is not present when simulating faults on the non-steered wheels.

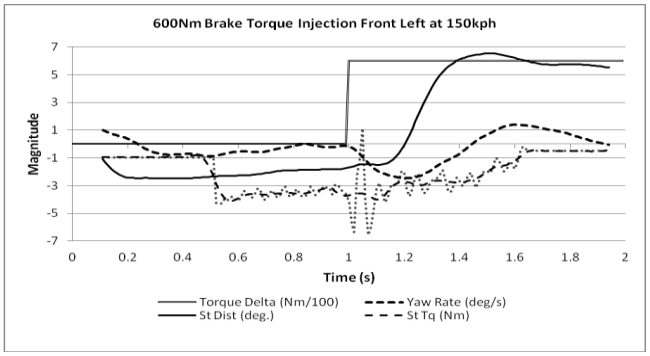

Figure 2: EV1 dynamic response to torque fault injections

The reaction times have been estimated for the 89 trial runs performed. The distribution is shown in Figure 3, which suggests that times are within generally accepted bounds i.e. less than 0.5s for a reflex action.

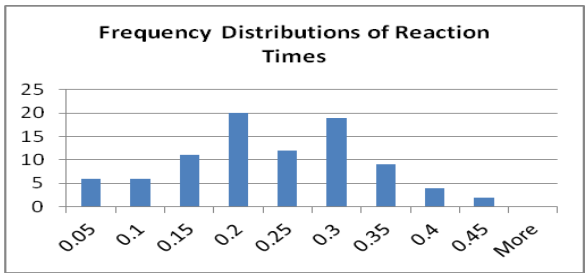

Figure 3: driver reaction time frequency distribution

Also of interest is whether there is any significant difference in reaction times for different levels of disturbance. This is plotted in Figure 4, where times are plotted for the front and rear axles; with values of less than 0.1s being excluded. From the graph there appears to be no significant difference between the values reported. If we assume that both data sets follow a normal distribution then the Student's t-test supports the null hypothesis that there is no difference.

Finally we can compare how the observed yaw rates compare with the limits given in Neukum [18]. Included in Figure 5 are simulation results for a generic vehicle having the same dimensions, tyres and weight (2,300kg) as EV1 (red), along with a trend line fitted to the data collected for the vehicle (blue). It can be clearly seen that at 600Nm the vehicle is at the defined limits.

4.5 The effect the vehicle has on controllability

When considering the driver's ability to control a given hazardous scenario we tend to think solely about the driver's ability to control the vehicle in the given scenario; and not consider the influence the vehicle's dynamic behaviour may have that on the driver's task.

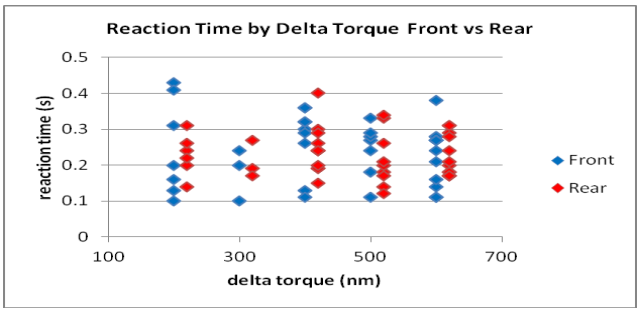

Figure 4: driver reaction time to torque faults of different magnitude and position

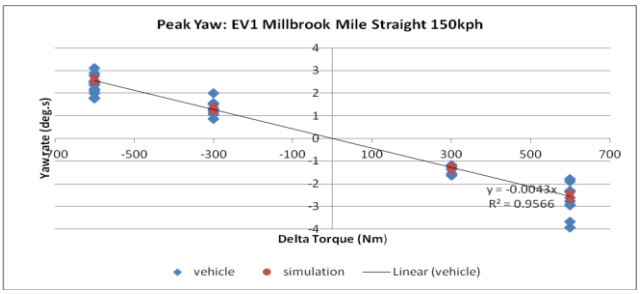

Figure 5: yaw rate data for EV1, vehicle observations compared with simulated results.

During normal driving the action of the driver can be thought of largely as a command and control task [8]. In order to achieve the command task the driver looks out of the window and uses steering inputs to keep the vehicle pointing in the direction they wish to go, and brake and accelerator inputs to maintain the desired speed.

For the average driver the control task associated with the above command task is likely to maintain the vehicle in the linear region of its response. That is, turning the steering wheel twice as much will result in the vehicle's radius of turn being twice as small. Less obvious to the driver as they control their vehicle through a bend is the vehicle's sideslip or yaw response. This is the vehicle's resonant behaviour in the ground plan and is analogous to the vehicle acting like a pendulum with the imaginary pivot point being ahead of the vehicle. The weight distribution of the vehicle (50:50 split, tail heavy, nose heavy) affects the level of side slip damping and consequently the vehicle's sensitivity to speed.

Large steering inputs, perhaps resulting from the driver taking mitigating action, may lead to the vehicle operating outside the linear response region. The biggest impact of this non-linearity on the driver's control task is that the vehicle may no longer be pointing broadly in the direction of travel; a situation that the average driver is known to be incapable of dealing with. This is borne out by the fact that brake based electronic stability protection (ESP) is mandated in the majority of territories. Again weight distribution, tyre cornering stiffness, and suspension tuning all affect the yaw damping ratio and the driver's ability to control the vehicle under such conditions.

4.6 What is dangerous?

As automotive functional safety engineers we readily use phrases like controllability, controllable by the average driver, difficult to control, and dangerous – this paper is no exception. But, what do we really mean? And when attempting to run experiments that quantitatively assess controllability, how should we define and interpret 'dangerous'? Should limits be set for the point at which the given scenario really becomes dangerous, or where it just feels dangerous?

This question has challenged the authors during the development of the in-wheel motor safety concept. As described in section 4.4 above, a failure within an in-wheel motor can lead to an un-commanded yaw moment being induced in the vehicle, which from the driver's perspective would result in the vehicle failing to travel along the desired path. But when does this deviation from the desired path become dangerous?

Technically an accident would be unlikely to occur until the given vehicle has exited its own lane and hit a vehicle or other object outside the lane. However, the discussions above about the social-technical impact of technology (section 1.3) and the effect of the vehicle on controllability (section 4.5) both suggest that a failure that results in the vehicle exiting its own lane may have exceeded what would constitute 'dangerous'.

The above considerations led to the limit of unintended yaw being set to that point at which the vehicle reached the edge of its own lane [8]; a level of yaw moment which also aligned with the Neukum *et al*. study [18].

5 Discussions

In general, for power train components the exposure and severity ratings will be fixed for any given situation. Thus, the only parameter in the risk matrix that we can significantly influence is controllability. Unfortunately, because we cannot pre-ordain the driver's response to a particular situation, we have no direct influence over how much control the driver or other persons will actually apply in any particular situation. Therefore, we require a good understanding of how our actors will respond to the failures, and the resulting hazard, to have a realistic idea of controllability for any given situation.

As will be evident from the preceding discussion, it is infeasible to expect all, or even a significant proportion of drivers, to respond in complex ways to emergency situations. Given this information, if the control of any given situation requires the driver (or other participants) to make complex inputs, then the level of stress and urgency associated with an event needs to be kept correspondingly low. Unfortunately, this may not be the case with events that significantly affect the vehicle dynamics.

Unsurprisingly (in hindsight) this has been known for some time in the human factors community. Dilich *et al*. [5] states that "*once it is determined that a driver was confronted with a sudden emergency demanding extraordinary response, the outcome of the accident is dictated more by the chance of the circumstances than by the performance abilities of the driver and his vehicle*".

5.1 First heresy: unintended correlation

A nagging question remains; why does the allocation of severity appear to be such a complex issue? From the perspective of an individual accident, an assignment of S3 seems hard to escape for all but the very lowest speeds. However, when examined from the perspective of large sets of data, S3 as defined in the standard appears to be virtually impossible. Instinctively both views can't be right, but is that actually true?

Considering accident statistics as a conglomeration of all available data, grouping similar accidents in order to construct a standalone E4 class of accidents (e.g. the way the NHTSA accident topology study [20] was performed), leads to a low expected severity. However, as we slice and dice the data, it is possible that as the exposure goes down, the severity for the selected subsets rises.

This is reflected in data derived from studies of specific accidents types (e.g. those involving foot traffic) in specific locations. These contain detailed information on injuries for a small subset of the available data, but have little or no information on occurrence rates; i.e. exposure, for the population as a whole as discussed in [27] which is critical of much early work.

Other confounding factors exist, such as the generally poor mapping between the types of data set used in accident investigations; with data collected by police (STATS19) and data collected by medical professionals generally being assessed on MAIS scale [16] [27] showing marked differences.

Another example is the use of delta V as a surrogate for impact severity. While it is useful to rate individual accidents using this metric, it is usually assessed *after the fact* and can have little correspondence with the information available for the generic scenario; such as the posted speed limit. Farmer [6] noted that at speeds greater than 80 kph, 87% of vehicles had a delta V of less than 40 kph and at 96 kph is was still 79%.

A possible compounding factor here is a human propensity to concentrate on the worst case scenario that could be conceived. Thus we find ourselves on the seesaw shown in Figure 6.

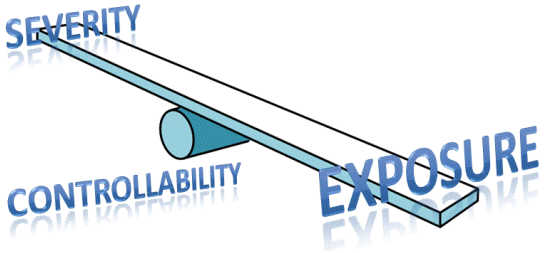

Figure 6: the severity/exposure seesaw

5.2 Second heresy: confounding control

In addition to the problem of uniquely assigning severity, based on either detailed studies or on general accident statistics, we have the issue of controllability as a confounding factor. In a high percentage of accidents the driver and other persons involved probably had some level of control; an observation supported by Farmers results [6]. The exact amount of control is another matter, but other than the driver, we generally have little idea about the level of control exercised by *all* other participants, although "flocking" behaviour may be significant [28]. Unfortunately, the concept of partial control is missing from ISO 26262. As indicated by [1], drivers often take *some* action which mitigates the severity of an accident even if they do not take enough to avoid the accident completely.

If partial control were to be considered then the worst case interpretation of severity would perhaps be the best fit. Otherwise, in the absences of the partial control concept, it may be better to rely on accident statistics, sliced and diced to suit the purposes of ISO 26262 and converted to a suitable scale in the manner of [9] [17].

5.3 Twenty questions

The MISRA Guidelines on controllability list a number of factors that can be taken into account when assessing controllability: human reaction times, ease of recognition of a situation, attentiveness, driver experience, smooth and readily perceived transfer of control from a system to the driver, and driver workload [15].

The process we used to satisfy ourselves about in-wheel motor controllability limits suggests that it may be possible to codify *some* of the controllability rating determination process. As an example, some of the questions that need to be addressed are listed below.

Question: is the failure under consideration dealt with by a normal driving response or is an unusual response needed?

Example: a 'side to side' brake performance discrepancy of less than 25% is deemed the acceptable limit for a vehicle to pass its Ministry of Transport (MOT) test [23] suggesting that a limited amount of yaw and necessary driver correction is "normal".

Question: will the driver's emergency response require cognition?

Example: steering and braking are examples of emergency responses that may not require cognition. However, placing the vehicle in neutral while it is moving requires cognition.

Question: will the primary effect of failure be obvious to the driver?

Example: an in-wheel motor failure that generates a significant torque disturbance would induce a strong yaw moment. However, a small discrepancy would be covered by normal driving.

Question: which senses will the driver use to detection the presence of the failure?

Example: a large in-wheel motor failure will be readily detected by the vestibular system and kinaesthetic senses. There may be hepatic feedback (torque steer) if the failure occurs on the steered (i.e. front) wheels.

Question: will an appropriate *automatic* response be available to the driver?

Example: the sudden glare of brake lights as noted by Green [7], yaw moment from steering failure [18], or in-wheel motor failures that mimic wind gusts [24], all lead to automatic driver responses.

Question: will the driver's required control actions that fall within normal bounds be sufficient?

Counter Example: we know that drivers may not use the vehicle brakes to their full physical capabilities [1], [22].

Question: are the safety goals (i.e. the high level safety requirements) compatible with the human factors?

The last question is really a meta-question, i.e. to review the decisions made.

6 Conclusions

When considering the development of safe automotive systems, the final point from the previous section is critical for two reasons: Firstly, if the safety goals do not complement the driver's behaviour then the requirements must be wrong; as one cannot change the driver! Secondly, if the safety goals to which the "item" is being developed are incorrect then it doesn't matter if you apply design rigour consummate with ASIL D, the system is potentially unsafe.

The difficulties associated with correctly identifying the needs of the end user, and effectively communicating requirements from domain users and experts to software experts is discussed by Leveson [13] in the context of autopilot development; with the observation being made "*that most errors found in operational software can be traced to requirement flaws, particularly incompleteness.*"

Within this paper we have discussed issues with determining severity, probability of exposure, and controllability of a given hazard situation. We have seen that although there are complications associated with determining values for severity and exposure, but for any given scenario they are largely fixed. This leaves controllability as the only factor over which the developers of automotive systems may have any influence.

The human factors literature tells us that if the action required by the driver is largely automatic then a fast driver response can be expected. However, the automatic and fast action taken by the driver may lead to an undesirable outcome if the driving "cues" received by the driver's sensory inputs trigger the "wrong" learned response. For example, the first time an inexperienced driver experiences "under-steer", resulting

from the front wheels loosing grip on an icy road, their learned response maybe to increase the hand wheel angle which will exacerbate the situation.

The take-home message from this work has to be that if you don't understand how the driver is going to react to a given situation, you can't be sure that you have captured the safety requirements correctly to ensure vehicle controllability is maintained. Consequently, it doesn't matter how rigorous you've been with your design (that is, by assigning the correct ASIL rating), if the requirements do not match the way in which the driver will respond then the system is still unsafe, even if the design rigour is perfect.

Acknowledgements

Many thanks to Damian Harty of Coventry University for the photo in Figure 1, for many useful thought provoking discussions and for commenting on this paper.

References

[1] Brake Technology Handbook, Ed. B. Breuer, K.H. Bill, SAE 2008.

[2] Crawford L.E, Cacioppo J.T. "Learning where to look for danger: integrating affective and spatial information", Psychol Sci. 2002 Sep;13(5):449-53.

[3] Department of Transport, "Reported Road Casualties in Great Britain: 2010 Annual Report, Contributory factors to reported road accidents".

[4] Dewar, R. Olson, P, "Human Factors in Traffic safety"

[5] Dilich, M., Kopernik, D., and Goebelbecker, J., "Evaluating Driver Response to a Sudden Emergency: Issues of Expectancy, Emotional Arousal and Uncertainty," SAE Technical Paper 2002-01-0089, 2002.

[6] Farmer C. M, "Reliability of Police-Reported Information for Determining Crash and Injury Severity", Traffic Injury Prevention, 4:1, 38-44, 2003.

[7] Green M., "How long does it take to stop?: Methodical Analysis of Driver Perception-Brake Times" in "Forensic Vision With Application to Highway Safety" 3rd Edition,. Ed. M.Green, M.J. Allen, B.S. Abrams, L.Weintraub, Lawyers and Judges. 2008

[8] Harty D, Gada T, Blundell M, Ellims M, Monkhouse H.E, "In-Wheel Motors - Some ISO26262 Safety Considerations" Vehicle System Dynamics, to appear.

[9] Hight, P., Wheeler, J., Reust, T., and Birch, N., "The Effects of Right Side Water Drag on Vehicle Dynamics and Accident Causation," SAE Technical Paper 900105, 1990.

[10] ISO 26262 Part 1, Road Vehicles – Functional Safety, Part 1 Vocabulary, ISO 26262-1:2011(E), 2011.

[11] Johnson, A (2001) 'Unpacking reliability: The success of Robert Bosch, GmbH in constructing antilock braking systems as reliable products', History and Technology, 17: 3, 249-270

[12] Leach, J. "Survival Psychology", MacMillan Press 1994

[13] Leveson N.G., Engineering a Safer World: System Thinking Applied to Safety, The MIT Press, 2011.

[14] MacAdam C.C, Gleason M, Pointer J.D, Sayers M.W, "Crosswind sensitivity of passenger cars and the influence of chassis and aerodynamic properties on driver preferences" Vehicle System Dynamics, 19(4), 1990, p. 201-236.

[15] The Motor Industry Software Reliability Association (MISRA), Development Guidelines for Vehicle Based Software, November 1994.

[16] Morris A, Mackay M, Wodzin E, Barnes J, "Some Injury Scaling Issues in UK Crash Research", Proc. Ircobi Conf., Lisbon, Portugal Sept. pp. 283–292 (2003).

[17] Najm W.G, Smith J.D, and Yanagisawa M, "Pre-Crash Scenario Typology for Crash Avoidance Research", DOT HS 810 767, April 2007.

[18] Neukum A, Ufer E, Paulig J, Kruger H. P, "Controllability of Superposition Steering System Failures", Steering Tech 2008, Munchen.

[19] Neukum A, "Controllability of erroneous steering torque interventions: Driver reactions and influencing factors", Steering Tech 2010.

[20] Pre-Crash Scenario Typology for Crash Avoidance Research, US Department of Transport, National Highway Traffic Safety Administration, DOT HS 810 767, April 2007.

[21] Staal M.A, "Stress, Cognition, and Human Performance: A Literature Review and Conceptual Framework."

[22] Verma, M. and Goertz, A., "Preliminary Evaluation of Pre-Crash Safety System Effectiveness," SAE Technical Paper 2010-01-1042, 2010.

[23] VOSA, "The MOT Inspection Manual: Private Passenger and Light Commercial Vehicle Testing" Fourth Edition, 2011.

[24] Wierwille W.W, Casali J.G, Repa B.S, "Driver Steering Reaction Time to Abrupt-Onset Crosswinds, as Measured in a Moving-Base Driving Simulator", Human Factors, 1983, 25(1), pg. 103-116.

[25] Young, M.S. and Stanton, N.A. "Back to the future: Brake reaction times for manual and automated vehicles", Ergonomics, 2007 50(1),46-58, 2007.

[26] J. Broughton, M. Keigan, G. Yannis, P. Evgenikos, A. Chaziris, E. Papadimitriou, N.M. Bos, S. Hoeglinger, K. Pérez, E. Amoros, P. Holló, J. Tecl, Estimation of the real number of road casualties in Europe, Safety Science,2010 Mar 48(3) 365-371.

[27] Erik Rosén, Helena Stigsonb, Ulrich Sandera, "Literature review of pedestrian fatality risk as a function of car impact speed", Accident Analysis and Prevention. 2011 Jan;43(1):25-33.

[28] Helbing, D. "Traffic and Related Self-Driven Many-Particle Systems", Rev. Mod. Phys. 73(4), 1067–1141 (2001)