1 Introduction

Human factors scholars have long recognized that circumscribing the human role in a human-machine system does not necessarily make that system less susceptible to human failure.\(^1\) This “iron[y] of automation”\(^2\) is also present in tort law: Injuries associated with automated vehicles and other robots will result at least in part from the behavior of the humans who use or encounter these machines. To mix metaphors, the human in the loop may be the weakest link in the chain. For this reason, a developer of an automated system must understand human behavior at least as well as it understands machine performance.

This paper uses the example of automated motor vehicles to illustrate the importance of the human user to the definition of a robotic system, to the safe design of that system, and to the analysis of the product liability exposure posed by that system. Its identification of three key yet uncertain targets—legal, safe, and actual use—offers a basis for broader analysis of human factors in robotic torts. Its specific discussion on a seller’s duty to warn in an evolving technological, business, and legal environment addresses an issue of likely importance in future litigation.

2 Robots as Joint Human-Machine Systems

The Defense Science Board does not mince words. “The pervasive effort to define autonomy and to create vehicle autonomy roadmaps,” it wrote, “is counterproductive”

---


\(^2\) Id.
and “a waste of both time and money.” Indeed, defining “levels of autonomy … deflects focus from the fact that all autonomous systems are joint human-machine cognitive systems, thus resulting in brittle designs.” The board concluded its two-page critique by noting that “all systems are supervised by humans to some degree, and the best capabilities result from the coordination and collaboration of humans and machines.”

The board is likely correct that humans will play some role in the design, dispatch, or real-time direction of most robots for the foreseeable future. Consider, for example, the interplay between human and machine that is implicit in ISO 26262, an international automotive standard for approaching safety-relevant electric and electronic systems. Some of the active safety systems that fall within its scope, such as antilock brakes and electronic stability control, are designed to compensate for the shortcomings of the human driver. Conversely, the required integrity level of such a system depends in part on the driver’s ability to detect that the system has failed and on her ability to avoid the specific harm caused by that failure.

Increasing vehicle automation actually amplifies the complexities of this human-machine relationship. Notwithstanding the Defense Science Board’s critique of the general enterprise, the SAE On-Road Automated Vehicle Standards Committee, on which I serve, sought to illustrate some of the key issues in this relationship. Drawing on similar work by the German Federal Highway Research Institute (BASt), the US National Highway Safety Administration (NHTSA), this expert group flattened the potential universe of relevant “joint human-machine cognitive systems” onto a single taxonomical spectrum of road vehicle automation.

The resulting draft definitions document describes driving as a task that is shared, whether consecutively or concurrently, by human and machine. The machine, in this case, is a “vehicle automation system,” a term selected over the more generic “system” precisely because transportation is a world of systems within systems within systems, many of which have human elements. The draft identifies six levels of vehicle automation, each of which can be uniquely described by five variables:

- Whether steering, acceleration, and deceleration are executed by a human or by the automation system.

---

4 Id.
5 Id.
6 Technical experts are discussing application of ISO 26262 to automated vehicles, which the current version does not expressly contemplate.
7 ISO 26262-3 annex B
8 This is the only nonbinary variable: At the “assisted” level of automation, the human and the automation system share these functions.
• Whether the driving and traffic environment is monitored by a human or by the automation system.

• Whether the automation system affords sufficient transition time for a human to effectively resume input.

• Whether a human must be available while the automation system is engaged—which depends on whether the automation system reverts to a “minimal risk state” in the absence of human input.\(^9\)

• Whether the automation system is functional in all driving environments or only in certain driving environments.\(^10\)

To these important criteria I would add one more: Whether a human or the automation system has ultimate authority over steering, acceleration, and deceleration. This question of authority arises in a variety of circumstances, including when an automation system can respond quicker than a human,\(^11\) when the automation system acts in concert with others,\(^12\) and when deliberate human signals to steer, accelerate, or decelerate are inconsistent with the environmental constraints perceived by the automation system.\(^13\) The committee discussed this criterion at length but reasonably considered it beyond the scope of its initial taxonomy.

These definitions do not imply the actual abilities of today’s machines or of today’s humans. Instead, they merely delineate minimum respective roles for the vehicle automation system and the human driver for largely prospective levels of automation.

Similarly, as the draft expressly states, the levels of automation make no assumptions about legality. For example, even though the automation system is tasked with monitoring the roadway, the human driver might still be legally obligated to remain vigilant.\(^14\) The draft accordingly avoids the words “control” and “responsibility”: Control is a vague term susceptible to conflicting technical, legal, and popular connotations,\(^15\) and the meaning of “responsibility” is dependent on whether the sphere is moral, criminal, or civil and on whether the mood is prospective or retrospective.

While avoiding these terms for the same reason, it ventures into three equally challenging concepts: legality, safety, and actuality.

---

\(^9\) The SAE definitions treat these two questions as separate variables.

\(^10\) This variable imperfectly maps a multidimensional taxonomy onto a single axis.

\(^11\) An automated emergency intervention system (AEIS), for example, might brake or steer to avoid an imminent crash.

\(^12\) Vehicle platoons and automated intersections, for example, might require such active coordination.

\(^13\) If an AEIS brakes or steers, for example, a human might attempt to accelerate or countersteer.

\(^14\) Automated Vehicles Are Probably Legal in the United States

\(^15\) Automated Vehicles Are Probably Legal in the United States
3 Three Keys: Legal, Safe, and Actual Use

The field of human factors is concerned with the behavior of a product’s human users. In principle, contemplated product uses should be legal, safe, and intuitive. Legality asks whether the contemplated behavior is proscribed by law; a lawful use can still give rise to civil liability, whether through negligence, strict liability, or no-fault insurance. Safety asks whether that behavior involves “unreasonable risk,” a concept that is common to engineering, to administrative rulemaking, and to products liability and that frequently invokes some notion of cost-benefit analysis. Actuality asks whether the contemplated behavior differs from the actual (intuitive and attainable) behavior, that is, whether and how the product is misused or abused.

Tensions among the three key design targets suggest particular structural failures. A mismatch between legality and safety implies that law as written is inefficient because it is either too permissive or too restrictive. A mismatch between safety and actuality suggests that users are either uninformed or irrational. And a mismatch between actuality and legality suggests that law is either underenforced or obsolete. Texting-while-driving restrictions offer a compelling illustration of several of these failures: Drivers who text while behind the wheel are behaving illegally and unsafely but not unusually, and some bans on the practice may actually decrease road safety.

---

16 ISO 26262-1, 1.101
17 E.g., ISO 26262-1, 1.129 (defining “unreasonable risk” as “risk judged to be unacceptable in a certain context according to valid societal moral concepts”).
As I have argued elsewhere, the key problem in managing the liability risk posed by increasing vehicle automation is actually uncertainty. Indeed, there is uncertainty about how existing laws will apply to automated driving, uncertainty about what is reasonably safe, and uncertainty about how drivers will adjust to automated systems in the near and long terms. Before turning to some of these issues, I consider these questions in a related context.

3.1 The Example of Traffic Speed

Traffic speed provides just one illustration of the complex interactions among and within these three elements. Imagine that you are driving your car at 70 miles per hour; are you speeding? Answering this question is easy—so easy that it has numerous answers.\textsuperscript{20} Consider a few:

Legal speed.\textsuperscript{21} Speed restrictions have multiple sources within law:

- Statutory speeds. Many state vehicle codes specify maximum speeds for certain road classes, such as Interstate highways, rural two-lane roads, and urban streets.
- Posted regulatory speed. A road’s posted speed may be lower than its statutory speed. The state or local maintaining authority typically sets such a limit by reference to the road’s operating speed.
- Posted advisory speed. On some road segments, the maintaining authority may advise a speed that is lower than that which is posted. Enforcement of such an advisory speed is at most indirect, through the basic speed law.
- Basic speed. The “basic speed law” common to most states requires drivers to maintain a reasonable and prudent speed. If visibility is poor, for example, driving slower than the posted speed might nonetheless violate this requirement.\textsuperscript{22}

Even these limits are not absolute. Exceeding the statutory speed may be expressly authorized for emergency vehicles\textsuperscript{23} and, in a few states, for passing another vehicle.\textsuperscript{24}

\textsuperscript{20} The most common of which may be: “If you are behind me, yes; if you are in front of me, you are not going fast enough.”


\textsuperscript{22} Automated Vehicles Are Probably Legal in the United States

\textsuperscript{23} http://www.health.ny.gov/professionals/ems/policy/00-13.htm

\textsuperscript{24} http://www.mit.edu/~jfc/laws.html#tolerance
More ambiguously, certain emergencies may occasionally justify noncompliance. Furthermore, several states effectively preclude a speeding conviction based on a recorded speed that is near the posted speed.

Operating speed. Historically, traffic engineers have used actual traffic speed as an indirect measure of a road’s appropriate speed. The dominant metric is the 85th percentile speed, which is the speed that 85 percent of unimpeded vehicles are traveling at or below. A posted limit is typically set near this point (provided that it does not exceed the statutory limit), which in one sense implies that exactly 15 percent of drivers are speeding. This approach, however, has at least two significant flaws. First, the collective judgment of drivers may marginalize others who use or are impacted by the road, particularly pedestrians, cyclists, and neighbors. Second, that judgment persistently underestimates certain risks.

Uniform speed. The speed of other vehicles may be as relevant to other drivers as to traffic engineers. Smoother traffic tends to be safer traffic—although this effect is not consistently observed in the context of differential speed limits for trucks and cars.

Efficient speed. For an individual, the optimal travel speed minimizes her marginal travel costs related to time, fuel, and personal crash risk. This individually optimal speed likely differs from the socially optimal speed, which would also account for impacts on the natural and human environments, health and safety, roadway capacity, and public resources. Notably, for example, Congress indirectly imposed a national maximum speed limit in 1974 primarily to reduce energy consumption and retained it in some form for over 20 years in large part because of the perceived safety benefits.

Survivable speed. Ralph Nader’s seminal book on the automotive industry of the mid 20th Century sought in part to rebut the broad perception that high-speed vehicle crashes were simply unsurvivable. Since then, improvements to active and passive safety systems in vehicles have contributed to a significant decline in crash-related fatalities and injuries. Nonetheless, every year in the US alone, motor vehicle crashes kill more than 30,000 people and injure more than two million. A pedestrian is highly

---

29 Managing Autonomous Travel Demand
30 http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_w90.pdf
31 http://www.census.gov/compendia/statab/2012/tables/12s1106.pdf
unlikely to die if struck by a vehicle traveling at 20 mph and highly likely to die if struck be a vehicle traveling at twice that speed; every 1 percent change in speed is associated with an increase of about 4 percent in fatal crashes.\textsuperscript{32}

Design speed. Transportation engineers frequently use a particular “design speed” to determine minimum geometric requirements for a road, including how sharply it can turn and how far along it a driver must be able to see. Only recently have state authorities begun to separate design speed from the concept of safe speed.\textsuperscript{33} Inherent in the design speed, which often exceeds the posted speed (whether regulatory or advisory), are myriad assumptions about environmental conditions and vehicle capabilities: A truck on wet pavement will handle differently than a sports car on dry pavement, for example. Furthermore, design speed has, at least prior to modern traffic calming approaches, functioned as a floor but not a ceiling: A horizontal curve might be widened or banked to accommodate travel at the design speed, but extraneous curves would not be introduced solely to discourage higher speeds. As one state department of transportation noted, this could yield “an infinite Design Speed” on a straight and flat road.\textsuperscript{34} Such a speed would be difficult to exceed.

In other words, legal speed affects design speed, which affects safe speed, which affects actual speed, which affects legal speed, which affects actual speed, which affects safe speed. These three concepts are distinct and dynamic, and designing a road that satisfies all of them without distorting any of them may well be impossible.

### 3.2 Application to Levels of Automation

Return to the SAE committee’s draft definitions from above. Of the six specified levels of automation, two in the middle provide particularly useful illustrations of the potential discrepancies among legal, safe, and intuitive use.

At a “partial” level of automation, the human operator monitors the driving and traffic environment and steers, accelerates, and decelerates only when necessary. Within this category, the safe and the legal human behaviors are largely coterminous; only New York expressly requires the human driver to keep at least one hand on the steering mechanism at all times. The more pressing question is whether this legal and definitionally safe behavior is actually attainable: Would a human who is not actively providing input to her vehicle be willing or even able to maintain the level of vigilance that is assumed?

\textsuperscript{33} http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_504.pdf at 7-8.
\textsuperscript{34} http://www.state.nj.us/transportation/eng/documents/RDM
At a “conditional” level of automation, the vehicle automation system monitors the driving and traffic environment, and the human operator steers, accelerates, and decelerates only when prompted in advance to do so by the vehicle automation system. Within this category, the safe and the legal human behaviors are, arguably, still largely coterminous; while many states require due care and prohibit distracted driving, most of these provisions specify or imply that a driver is only distracted when she is behaving unsafely.

Again, the more pressing question here is whether this legal and definitionally safe behavior is actually attainable: Would a human who is not even monitoring the roadway be willing to stay in an optimal driving position and able to stay awake? Would she be willing and able to acquire and then to retain the driving skills necessary to safely maneuver the vehicle when actually required? Moreover, when her vehicle encountered unusual road or traffic conditions, would she be willing to defer to the automation system’s response?

3.3 The Real-Time Human Role in Safe Operation

These questions implicate the very idea of safety as well as the values that underlie or conflict with it. As the discussion of travel speed indicated, there is no uniform understanding of safety from a technical much less legal perspective. Indeed, safety assessments necessarily involve assumptions about scale, timeline, and causation—classic issues related to system boundaries.

Consider, for example, how well a “safe” human-vehicle pair must be. One answer might be that it must perform at least as well as an expert human driver for any conceivable individual maneuver or scenario. Such a stringent standard, however, might mean that automation technologies reach the market more slowly and at greater cost, resulting in the loss of lives that they could have saved. The dramatization: Perpetuation of a tragic status quo.

An answer on the other side of the spectrum might be that the human-vehicle pair must, in a broad statistical sense, be safer than today’s cars and drivers. Based on rough 2009 estimates, a vehicle is involved in a crash about once every 160 thousand miles and in a fatal crash about once every 65 million miles. Automated vehicles that could simply beat these figures might, in broad terms, represent a safety improvement. Because crashes involving such vehicles might be different than those exclusively involving more conventional vehicles, however, the result of this logic might be fatalities that would be preventable under today’s assumption of unambiguously human-driven vehicles. The dramatization: Headlines proclaiming “machines kills child,” ruinous litigation, and long-term damage to the credibility of the technologies.

---

35 http://cyberlaw.stanford.edu/blog/2012/03/driving-perfection
Related to this question is the role of the human. Conceivably, for example, an automated vehicle might not require real-time human monitoring to satisfy a particular safety threshold. Nonetheless, what if that human oversight further increased the system’s overall safety? Alternately, what if occasional human intervention were to prevent some crashes while causing others? For the reasons discussed in an earlier work, deliberately eclipsing human authority—the variable excluded from SAE’s initial draft—may well alienate actual use from legal use even if it more closely aligns actual and safe use.

On the assumption that humans may still need to maneuver their vehicles in some situations, particularly emergencies and unusual conditions, degradation of or failure to acquire critical driving skills is a concern. Unclear, however, is how that concern should factor into an assessment of safety. A recent safety alert from the Federal Aviation Administration (FAA) raises this point: While acknowledging that autoflight systems “have improved safety,” the memorandum nonetheless urges airlines to ensure sufficient opportunity for “manual flight operations” to prevent “degradation of the pilot’s ability to quickly recover the aircraft from an undesired state” 36—a contributing factor to, for example, the 2009 crash of an Air France flight. 37 But while airline passengers might understand the need for practice, they might not appreciate being on the flight used for that practice.

4 Dynamic and Custom Warnings

4.1 The Manufacturer-Vehicle and Manufacturer-User Relationships

Although early vehicles appear to have been sold both directly by their manufacturers and indirectly through dealers, 38 the dealership model eventually became dominant and, by the 1950s, codified in law. 39 At that time it was argued that, in contrast with the large manufacturers, “local dealers might have greater sensitivity to local consumer preferences and provide better overall service.” 40

40 http://www.ftc.gov/speeches/leary/learystateautodealer.shtm
Any lack of privity between manufacturers and the actual users of their products was of questionable relevance to early automotive litigation and became even less important with the advent of modern product liability law. Under various theories that emerged or coalesced, automakers were and remain liable for foreseeable injuries caused by defects in their vehicles.

The debut of GM’s OnStar in the mid-1990s marked a further integration of vehicle makers and users. This subscription service, and similar offerings from other automakers, provide a direct link to and from a customer’s vehicle. The potential for data collection is huge—and not necessarily contingent on a subscription or other ongoing contractual relationship. Many more cellular-based automotive telematics applications are likely in the near term, and an imminent regulatory decision on dedicated short-range communications (DSRC) could substantially expand vehicle connectivity in the medium to long term.

Tesla’s over-the-air update of its vehicle software in late 2012 stands as another watershed moment. This industry first means that vehicles themselves have become dynamic rather than static products, a development with implications for current legal approaches to vehicle certification, recall and improvement, and defect analysis. Because many of these issues address the performance of the vehicle independent of the human operator (if any), I do not consider them here.

As a result of these legal and technical innovations, automakers today probably have a closer relationship with—and know more about—the buyers, owners, and users of their products than at any time since the companies’ founders were personally selling cars.

Now extend this premise into a future of increasing but not total vehicle automation. Safe operation would still depend in part on the vehicle, its environment, and the performance of its operator, all of which could to some extent be knowable to the automaker through various telematics—in some cases more so than to the operator herself. That operator, meanwhile, may repeatedly engage in transactions with, receive

---

41 Graham at 1245-56.
43 http://digitalcommons.law.scu.edu/cgi/viewcontent.cgi?article=1556&context=chtlj
47 http://wired.com/autopia/2012/09/tesla-over-the-air. For a sample of some of the devices that support firmware over-the-air updates, see http://www.redbend.com/en/fota-enabled-device-gallery/gallerymenu
information\textsuperscript{48} from, or use features under terms specified by that automaker or one of its partners.

\subsection*{4.2 Current Duties to Warn}

The expanding literature on the application of product liability law to automated vehicles has tended to focus on issues other than an automaker’s duty to provide adequate warnings.\textsuperscript{49} The human factors issues involved in various levels of automation, however, suggest that this “continuously enlarging”\textsuperscript{50} duty will remain highly relevant. Even today, claims regarding this duty are “among the most common allegations in products liability litigation.”\textsuperscript{51}

They proceed from the principle that “a seller will have a duty to provide warnings as to the risks of use or consumption of a product where the risk is material and the seller knows or should know that the user is less informed concerning that risk than the seller.”\textsuperscript{52}

Depending on the jurisdiction, another duty to warn may exist—and extend beyond the date of the product’s sale. Unlike its predecessor (and, arguably, persistent competitor), the Restatement (Third) of Torts: Products Liability contemplates liability for a seller or

\textsuperscript{50} Madden at 234.
\textsuperscript{52} Madden at 234.
Section 10 states in part that:

(a) One engaged in the business of selling or otherwise distributing products is subject to liability for harm to persons or property caused by the seller’s failure to provide a warning after the time of sale or distribution of a product if a reasonable person in the seller’s position would provide such a warning.54

Although a uniform and consistent post-sale duty did and does not exist among the states,55 the Restatement (Third)’s provision has “received a mixed but generally favorable reception in the law since 1998,”56 and “more than thirty states have [since] adopted various versions of duties arising after the sale of a product.”57

In order to limit what its reporters described as “the most expansive area in the law of products liability,”58 the Restatement (Third) further provides that:

(b) A reasonable person in the seller’s position would provide a warning after the time of sale if:

(1) the seller knows or reasonably should know that the product poses a substantial risk of harm to persons or property; and

(2) those to whom a warning might be provided can be identified and can reasonably be assumed to be unaware of the risk of harm; and

(3) a warning can be effectively communicated to and acted on by those to whom a warning might be provided; and

(4) the risk of harm is sufficiently great to justify the burden of providing a warning.59

The technological developments sketched above could make several of these elements easier to achieve (and to establish) than section 10’s drafters might have imagined. The operators of a vehicle may be reached, if not personally identified, through that vehicle’s

---

54 Restatement (Third) of Torts: Product Liability sec. 10
55 Ross at 963-64.
56 Ross at 984.
57 Stilwell at 1036.
58 Ross at 965 (quoting Henderson and Twerski).
59 Restatement (Third) of Torts: Product Liability sec. 10.
telematics; the potential recipients of a warning could also conceivably include other vehicle occupants, other road users, and perhaps even law enforcement. And unless broader interests of privacy or consumer preference are considered, the actual burden to the manufacturer of supplying such a warning might be small in comparison to more traditional mailings or media campaigns.

Like the more conventional time-of-sale duty, the post-sale duty is premised in part on the information disparity between seller and user.60 This is, interestingly, a typical rationale for negligence rather than strict liability,61 although design and information defects tend to blur the distinction between the two.62 Other bases for a post-sale duty involve a close connection between the manufacturer and its products or customers.63

4.3 Speculative Manifestations of Duties to Warn

Vehicle automation will likely make interactions between vehicles and their operators more complex; facilitating the safety of these interactions will be significant design challenge that will rely in part on effective communication; manufacturers are increasingly connected to their vehicles as well as to the people who use them; and theories of liability premised on a seller’s continued connections with and knowledge about its products are expanding.

These observations suggest that warnings may be prominent in vehicles and that failure-to-warn claims may be prominent in courtrooms. Some of these actual or absent warnings may be conventional: Particular cautions about the limits and requirements of the vehicle automation system that are printed on the vehicle or in the owner’s manual. The more conceptually interesting, however, may push both technology and law.

One of the promises of automated vehicles lies in potential to ascertain and address risks better than a human driver. And yet the premise of this paper is the interdependence of machine and human in the operation of such vehicles. In what way, then, might a failure of the human user also constitute a failure of the product itself—particularly, when the vehicle or its manufacturer may well know more about the limitations of the human user than she herself? Signs of inattention, fatigue, and carelessness, for example, may be evident in the user’s inputs into the vehicle even if they are not manifest to the user herself.

Manufacturers or their vehicles might also be better positioned than these users to recognize specific contextual hazards of vehicle operation related to the vehicle and its environment. While the general risk posed by speeding may be obvious to the user and

60 Id. sec. 10(b)(1)-(2).
61 Stilwell at 1045.
62 Cf. Stilwell at 1045-46.
63 E.g., Stilwell at 1044-45; Ross at 970-71.
hence not subject to a duty to warn, for example, speeding under specific conditions may pose particular and arguably nonobvious risks. The vehicle automation system's level of confidence in its classification of a particular road object, for example, might likewise be relevant information for the user, and specific pavement and traffic condition data may also be more readily perceived by the vehicle than by its user. Numerous potential risks may be identifiable in the information silently collected or generated by the vehicle.

To the extent that these risks are addressed by more conventional printed warnings, an injured user—or, perhaps more sympathetically, an injured third party—might argue that more dynamic, specific, and contextual warnings would have been more effective. As more of the driving experience is customized, that plaintiff might argue, shouldn't warnings be as well?

The data collection potential of vehicles evokes two additional theories. As a basis for recovery, the implied warranty of fitness for a particular purpose is limited: It can generally be disclaimed and is not helpful to third-party plaintiffs. In a broad sense, however, it suggests that the flow of information from buyer to seller can increase the obligations of that seller with respect to its product. Somewhat similarly, the tort of negligent entrustment is premised on an individual's known or knowable propensity to dangerously use an instrumentality. In both of these cases, one actor's increased knowledge of another can lead to liability.

5 Conclusion

Human users present human designers with a key challenge: How to facilitate the safe and legal use of robotic products that remain joint human-machine systems. The alignment of safe and actual use, while iterative and in some ways self-defeating, depends in part on the effective communication of risks and limitations. In the case of automated vehicles, related advances in automotive technology may significantly expand the opportunity for such communication—and the duty of manufacturers to engage in it.