Testimony of Mary Cummings, PhD  
Director, Humans and Autonomy Laboratory  
Director, Duke Robotics  
Professor of Mechanical Engineering and Materials Science  
Professor of Electrical and Computer Engineering  
Duke University  

Hands Off: The Future of Self-Driving Cars  

March 15th, 2016

Good afternoon Chairman Thune, Ranking Member Nelson, and distinguished members of the committee. Thank you for the opportunity to appear before you to discuss issues related to the future of self-driving cars in the United States.

I am the director of Duke Robotics and the Duke University Humans and Autonomy Laboratory, which focuses on the multifaceted interactions of humans and autonomous systems in complex sociotechnical settings. I have conducted driving research and provided future technology recommendations to automotive manufacturers for more than a dozen years including Ford, Nissan, Toyota, and Google X1. I was the program manager for a $100 million Navy robotics helicopter that carries sensors very similar to those on self-driving cars. I am also currently conducting research for the National Science Foundation on the interaction of self-driving cars and pedestrians.

While I enthusiastically support the research, development, and testing of self-driving cars, as human limitations and the propensity for distraction are real threats on the road, I am decidedly less optimistic about what I perceive to be a rush to field systems that are absolutely not ready for widespread deployment, and certainly not ready for humans to be completely taken out of the driver's seat.

The development of self-driving car technologies has led to important advances in automotive safety including lane departure prevention and crash avoidance systems. While such advances are necessary stepping stones towards fully capable self-driving cars, going from automated lane changing or automated parking to a car that can autonomously execute safe control under all possible driving conditions is a huge leap that companies are not ready to make.

Here are a few scenarios that highlight limitations of current self-driving car technologies: The first is operation in bad weather including standing water on roadways, drizzling rain, sudden downpours, and snow. These limitations will be especially problematic when coupled with the inability of self-driving cars to follow a traffic policeman's gestures.

Another major problem with self-driving cars is their vulnerability to malevolent or even prankster intent. Self-driving car cyberphysical security issues are real, and will have to be addressed before any widespread deployment of this technology occurs. For example, it is relatively easy to spoof the GPS (Global Positioning System) of self-driving vehicles, which involves hacking into their systems and guiding them off course. Without proper security systems in place, it is feasible that people could commandeer self-driving vehicles (both in the air and on the ground) to do their bidding, which could be malicious or simply just for the thrill and sport of it.

And while such hacking represents a worst-case scenario, there are many other potentially disruptive problems to be considered. It is not uncommon in many parts of the country for people to drive with GPS jammers in their trunks to make sure no one knows where they are, which is very disruptive to other nearby cars relying on GPS. Additionally, recent research has shown that a $60 laser device can trick self-driving cars into seeing objects that aren’t there. Moreover, we know that people, including bicyclists, pedestrians and others drivers, could and will attempt to game self-driving cars, in effect trying to elicit or prevent various behaviors in attempts to get ahead of the cars or simply to have fun.

Lastly, privacy and control of personal data is also going to be a major point of contention. These cars carry cameras that look both in and outside the car, and will transmit these images and telemetry data in real time, including where you are going and your driving habits. Who has access to this data, whether it is secure, and whether it can be used for other commercial or government purposes has yet to be addressed.

So given that these and other issues need to be addressed before widespread deployment of these cars, but understanding that there are clear potential economic and safety advantages, how can we get there with minimal risk exposure for the American public? In my opinion, the self-driving car community is woefully deficient in its testing and evaluation programs (or at least in the dissemination of their test plans and data), with no leadership that notionally should be provided by NHTSA (National Highway Traffic Safety Administration). Google X has advertised that its cars have driven 2 million miles accident free, and while I applaud this achievement, New York taxi cabs drive two million miles in a day an a half. This 2 million mile assertion is indicative of a larger problem in robotics, especially in self-driving cars and drones, where demonstrations are substituted for rigorous testing.

RAND Corporation says that to verify self-driving cars are as safe as human drivers, 275 million miles must be driven fatality free. So that means we need a significantly accelerated self-driving testing program, but it is not simply good enough to let self-driving cars operate in California or southern Texas to accrue miles. NHTSA needs to provide leadership for a testing program that ensures that self-driving cars are rigorously tested for what engineers call the “corner cases”, which are the extreme conditions in which cars will operate. We know that many of the sensors on self-
driving cars are not reliable in good weather, in urban canyons, or places where the map databases are out of date. We know gesture recognition is a serious problem, especially in real world settings. We know humans will get in the back seat while they think their cars are on “autopilot”. We know people will try to hack into these systems.

Given self-driving cars’ heavy dependence on probabilistic reasoning and the sheer complexity of the driving domain, to paraphrase Donald Rumsfeld, there are many unknown unknowns that we will encounter with these systems. But there are many known knowns in self-driving cars that we are absolutely aware of that are not being addressed or tested (or test results published) in a principled and rigorous manner that would be expected in similar transportation settings. For example, the FAA (Federal Aviation Administration) has clear certification processes for aircraft software, and we would never let commercial aircraft execute automatic landings without verifiable test evidence, approved by the FAA. To this end, any certification of self-driving cars should not be possible until manufacturers provide greater transparency and disclose how they are testing their cars. Moreover, they should make such data publicly available for expert validation.

Because of the lack of safety evidence, I agree with California’s recent ruling that requires a human in the driver’s seat. However, while I generally support individual state governance on these issues, the complexity of the operation and testing of robotic self-driving cars necessitates strong leadership by NHTSA, which has generally been absent. But as I testified in front of this committee two years ago\(^2\), the US government cannot and has not maintained sufficient staffing in the number of people it needs who can understand, much less manage, complex systems such as self-driving cars. So it is not clear whether NHTSA or any other government agency can provide the leadership needed to ensure safety on American roads.

Let me reiterate that as a professor in the field of robotics and human interaction, I am wholeheartedly in support of the research and development of self-driving cars. But these systems will not be ready for fielding until we move away from superficial demonstrations to principled, evidenced-based tests and evaluations, including testing human/autonomous system interactions and sensor and system vulnerabilities in environmental extremes. To this end, in collaboration with private industry, NHTSA should be providing strong leadership and guidance in this space.

King Coal Highway

Model Public–Private Partnership

U.S. Energy Outlook: Positive Implications for Transportation

Automated Vehicle Technology: Research Areas to Follow

Driverless Cars: Promises and Pitfalls

Programming Funds for Research

Plus: Practical Findings from Cooperative Research Programs
With the move toward driverless cars, including automated driving assistance in the short term, the appropriate levels of shared authority and what the interaction should be between the human driver and the automation remain open questions. How robust driverless cars may be against system failures—including human failures—and operating in degraded sensor environments is unclear; more principled research and testing are needed.

Automation on board vehicles is inherently brittle and can account only for what it has been programmed to consider. Communication between a technically complex system and humans with a varying range of driving and attention management skills is difficult, because the driver must be appropriately informed about the state of the system, including its limitations, and will need to build appropriate trust—neither too much nor too little—in the capabilities of the automation.

Further complicating the problem is that automated systems can lead to boredom, which encourages distraction, as a significant body of research has demonstrated. The operator therefore may be unaware of the state of the vehicle—leading to mode confusion—and may not respond quickly and appropriately in an accident. Over time, the degradation of operator skills as a result of automation can reduce the ability to respond to emergent driving demands and will likely lead to risk homeostasis—the increased tolerance of risk—even in normal operations.

Tests and Design Considerations
These issues are well-known to the human systems engineering community, but it is unclear whether driverless car designers are considering these issues or whether manufacturers are conducting appropriate human-in-the-loop tests with representative members of the driving population. Until tests show that the vehicles account for these issues, driverless cars will not be safe for unrestricted access and use on U.S. roadways.

Moreover, significant sociotechnical considerations do not appear to be a concern in the push to introduce this technology on a wide scale. The utilitarian approach, quoted by many in the press, is that driverless cars eventually will kill people, but that this is acceptable because of the likely reduction in total deaths. Nonetheless, the likelihood of a reduction is not yet proved. The utilitarian approach demonstrates insensitivity to a deontological perspective—that is, to moral obligations—which causes many people to be uncomfortable about a significant shift of responsibility and accountability from humans to computers.
Driverless or Driver Optional?

Driverless car technologies in development include the ability to navigate roadways, change lanes, observe traffic signals, and avoid pedestrians, without human input. These technologies require Global Positioning System (GPS) information, internal navigation maps, outward-facing cameras, and possibly the use of laser and other range-finding systems—the specifics of the systems vary by company.

The first two of these technologies allow the vehicle to understand where it is in the world, where it should be going, and how to get there; the latter two allow the vehicle to track where it is on the road and where other vehicles, traffic indicators, and pedestrians are. The active cruise control (ACC) systems now in some vehicle models are early forms of this technology; this limited form of autonomy can serve as a forerunner to more advanced systems.

Although termed driverless, the vehicles are better classified as driver optional, particularly under the National Highway Traffic Safety Administration’s (NHTSA’s) Levels 2 and 3 of automated driving, in which human operators have either primary or secondary control responsibilities (1). Although these vehicles supposedly are capable of driving in any traffic situation without requiring the human driver to apply pressure to the pedals, shift, or steer, the driver still may choose to do so and may play a role in avoiding crashes.

In the distant future, the driver will not be needed; the current autonomous driving systems, however, require a human to be in the driver’s seat and allow and—in some cases, expect—the driver to assume control at specific points. This is the problem: as long as a human operator has some expectation of shared authority—whether primary or secondary—the design of the automation must ensure that the operator fully understands the capabilities and limitations of the vehicle, maintains full awareness of what the system is doing, and knows when intervention might be needed. Failure to do this may lead to a variety of automation- and human-induced crashes.

Interacting Weaknesses

Google’s driverless cars already have logged more than 300,000 miles, with two reported crashes (2). One occurred when the car was traveling under manual control on roads not previously mapped into its system (3). The actual causal chain is disputed, but the event illustrates the brittleness of automation—the car may not be able to handle uncertainty in its internal model, and this can be exacerbated by human error.

These problems are aggravated by an inherent human limitation known as neuromuscular lag—even when paying attention perfectly, a person experiences a lag of approximately one-half second between seeing a situation develop and taking a responsive action. Instances of human error like this are not the fault of the human alone but of the interaction between the human and the automation and...
the weaknesses of each—the human’s imperfect attention and execution of a response and the automation’s brittleness in perception and in generating a solution.

Although computing reliable accident statistics would be premature, if driverless cars could sustain this crash rate, they would be an improvement over teenage drivers. According to the Insurance Institute for Highway Safety, teenagers are three times more likely to have a crash than drivers age 20 or older.

**Lessons from Aviation**

The driverless car community can look to aviation for lessons learned from the introduction of automation to relieve pilot workload and—in theory—improve safety. Since the introduction of increasing automation in flight control and navigation systems in the mid-1970s, the accident rate in commercial jet operations has dropped from approximately 4 per million departures to 1.4 (+4).

Automation has been key in reducing this accident rate. Nevertheless, many accidents labeled as human error by the Federal Aviation Administration (FAA) and the National Transportation Safety Board would be better categorized as failures of human–automation interaction. These include the following examples:

- A faulty indicator light that appeared on final approach caused the 1972 crash of Eastern Airlines Flight 401. Distracted by the disagreement between the warning light and other gauges, the crew failed to notice that the autopilot had been disengaged accidentally. No alert or warning notified the pilots, who focused on the indicator problem and failed to notice that the aircraft was descending steadily into the Everglades.

- Air France 447, which crashed off the coast of Brazil in 2009, involved two failures: failure of the automation and a failure of the displays to present information to the operator. A clogged pressure sensor caused the autopilot system to act as if the altitude of the airplane was too low. The autopilot put the aircraft into an increasingly high climb, eventually triggering the stall warning alert. With the aircraft on autopilot, the pilot was distracted and was not fully engaged in monitoring the aircraft; this is a common occurrence. When the stall warning activated, the pilot was not aware of what was happening and made the worst of all possible decisions—he attempted to increase the aircraft’s climb angle, which worsened the stall and contributed to the crash.

- Northwest Flight 188 overshot Minneapolis, Minnesota, by roughly an hour in the fall of 2009 as a consequence of operator boredom and resultant distraction. With the aircraft on autopilot, both pilots became distracted by their conversation and failed to monitor the aircraft and its status. As they opened their laptops to obtain information to supplement their conversation, they misdialed a radio frequency change, missed at least one text message sent by air traffic control inquiring about their location, and only realized what was occurring when a flight attendant asked about the landing time. Luckily, the result was only a late landing; more severe consequences could have occurred.

**Attention and Distraction**

These issues are common to many other domains involving human interaction with automated systems and are well known to the human systems engineering and experimental psychology communities. In general, the research community agrees that human attention is a limited resource to be allocated, and that the human brain requires some level of stimulus to keep its attention and performance high.

Without this input, humans seek it elsewhere, leaving them susceptible to distraction by either endogenous or exogenous factors. Operators may miss important cues from the automation or from the environment—as in Eastern Flight 401; or they may see the cues but not have all of the information required to make a correct decision—as in Air France 447; or they may use their spare capacity to engage in distracting activities, leading to a loss in situational awareness—as in Northwest Flight 188. An operator also may enter a state of mode confusion and make decisions believing that the system is in a different state than it actually is.

Although these examples and research come from aviation, the role of a pilot monitoring an aircraft autopilot system is similar to that of the human driver in a driverless car. Recent research in human–automation interaction has expanded to automated driving systems and is showing the same
Studies on human interaction with automated systems have shown that human attention is limited and distraction common when automation is active.

Effects of Automation

Drivers in an autonomous or highly automated car were less attentive to the car while the automation was active, were more prone to distractions, especially to using cellular phones, and were slower to recognize critical issues and to react to emergency situations, for example, by braking.

In tests, automated systems used at lower average speeds and with greater separation between vehicles yielded benefits, but at the cost of poorer performance by humans in emergency situations (5, 6). In other words, when the automation needed assistance, the operator could not provide it and may have made the situation worse. The operator cannot be assumed to be always engaged, always informed, and always ready to intervene and make correct decisions when required by the automation or the situation. This applies to highly trained pilots of commercial airliners, as well as to the general driving population of the United States and other countries, who receive little to no formal training and assessment.

Technology Robustness

Much of the development of driverless cars is proprietary, and the exact capabilities of the technologies are not known. This prevents definitive statements about a specific vehicle, but not comments on the limitations of the technology overall or specific questions of concern. Google's autonomous car—generally regarded as the most advanced—relies on four major technologies: lidar, or light detection and ranging; a set of onboard cameras; GPS; and maps stored in the vehicle's onboard computer. The GPS signal tells the car where it is on the stored map and where its final destination is, and from this, the car determines its route. Cameras and lidar help the vehicle sense where it is on the road, where other vehicles are, and where to find and follow stop signs and streetlights.

Each of these systems is vulnerable in some way, and the extent of redundancy is not known, or whether the car will function correctly if any one of the four systems fails. If the GPS or maps fail, the car does not know where it is on its route and where it should be going. If the lidar fails, it may not be able to detect nearby vehicles, pedestrians, or other features. If the cameras fail, the vehicle may not be able to recognize a stop sign or the color of the traffic light. Also not clear is how much advance mapping and how often map updates are required to maintain an effective three-dimensional world model by which the onboard computer makes decisions. Moreover, GPS signals can be unreliable in urban canyons in which tall buildings, tunnels, and other forms of structural shielding cause a lost or degraded signal.

Flaws in the Systems

The security of GPS is questionable. Spoofing or mimicking a GPS signal to provide false location information, as well as jamming or forcibly denying a GPS signal, has been observed by the U.S. military (7, 8) and in civilian applications (9). An individual or group of individuals spoofing GPS signals in major metropolitan areas during rush hour, for example, could force cars off the road, into buildings, or off bridges, or could cause other damage.

Google's researchers admit that they have yet to master inclement weather and construction areas (2). Precipitation, fog, and dust create problems for lidar sensors, scattering or blocking the laser beams and interfering with the image detection capabilities of the camera. As a result, the vehicle is unable to sense the distance to other cars or to recognize stop signs, traffic lights, and pedestrians.

Other research has noted that the technology cannot currently handle construction signs, traffic directors—a task that requires sophisticated recognition of gestures—and other nonnormal driving conditions (2). A related question is how well the system can anticipate the actions of other drivers; avoiding a car calmly changing lanes is entirely different from anticipating the actions of a reckless and irrational driver. Previous research has shown that people are prone to distraction; any failures or degradations in a technology that requires monitoring by humans will increase the likelihood of a serious or fatal crash significantly.

Trust and Skill Degradation

How drivers adapt to the presence and performance of the automation is not a trivial issue. If the automa-
tion is perceived to be unreliable or not proficient, then the operator refuses to use the system, despite the potential benefits. When automation is perceived as proficient, however, operators rely more heavily on the technology and fail to use their own skills. This leads to a loss of skill and increases reliance on the automation, possibly leading back to mode confusion, as discussed earlier.

Skill degradation from overreliance on automation is a problem in aviation; FAA recently released a safety notice recommending that pilots fly more often in manual mode than with the autopilot. Risk homeostasis is another possible concern—drivers perceive the automation to be more capable and begin to accept more risk; this leads to increased distraction and overreliance on the automated system.

Research into ACC systems already has observed some of these concerns. The 2014 Jeep Grand Cherokee owner’s manual states that ACC “is a convenience system… not a substitute for active driving involvement,” and the BMW Technology Guide notes that “the system is not intended to serve as an autopilot” (10). Nevertheless, studies addressing public knowledge of the capabilities of ACC systems show that the public is not fully aware of the limitations of the technology and has a poorly-defined sense of when to trust the autonomy and when driving should be a manual operation. In a series of experiments, many drivers displayed riskier behavior when given the ability to use the limited autonomy of ACC systems, including the failure to shut off the automated systems when conditions were not suitable (5).

Providing appropriate feedback to the operator on the performance of the operator and of the automation is crucial to mitigate these problems, but designing a system for appropriate trust is a challenge (11). The automation should be capable of describing its performance and its limitations to the driver, who should then be able to learn how best to use the automation in the course of driving. The automation also should be able to sense when the human operator is performing poorly, or even dangerously, so that it can either support the driver or take control. The end result is more of a partnership—each side understanding and accounting for the abilities and limitations of the other.

Sociotechnical Considerations

A common argument in favor of inserting driverless car technology as soon as possible is that accidents and fatalities will be reduced dramatically. According to Google’s Sebastian Thrun, “more than 1.2 million lives are lost every year in road traffic accidents. We believe our technology has the potential to cut that number, perhaps by as much as half” (12). Although a logical argument in keeping with rational decision-making theory, such a utilitarian approach is not universally shared. A deontological approach could assert that machines should not be allowed to take the lives of humans under any circumstances—which is similar to one of the three laws of robotics drawn up by author Isaac Asimov.

A lower fatality rate is not a guarantee with autonomous cars, particularly at NHTSA Levels 2 and 3, but if the fatality rate is lower than that with human-operated vehicles, the killing of a human by a machine, even accidentally, will not resonate well with the general public. Recent intense media and public campaigns, for example, have protested autonomous weaponized military robots. Similar issues are likely to be raised if driverless or driver-assisted technology is responsible for a fatality or a serious accident that receives intense media attention.

Furthermore, the chain of legal responsibility for driverless or driver-assisted technologies is not clear, nor is the basic form of licensure that should be
required for operation. Manufacturers of driverless technologies and the related regulatory agencies are responsible not only for considering the technological ramifications but also the sociotechnical aspects, which have not been addressed satisfactorily.

Tenuous Transition
Driverless car technology promises potentially safer and more efficient driving systems, but many questions remain. The robustness of the technology and the interaction between the human driver and the driverless technology are unclear. Boredom and distraction, mode confusion, recovery from automation errors, skills degradation, and trust issues are major concerns and have been observed in experimental and real-life settings. Solutions to these problems will come through proper design, supplemented by extensive testing to confirm that the solutions are having the intended effect.

Manufacturers have not provided any documentation, including extensive, independent, and principle testing, describing how their designs have addressed these issues. Moreover, these issues lie outside the typical tests that regulatory agencies perform in assessing safety. Until these issues have been addressed through independent human-in-the-loop testing with representative user populations, these vehicles should remain experimental. Public- and private-sector organizations alike should develop testing programs, as well as programs to test the reliability and robustness of the core technologies such as GPS and lidar, and should set requirements for driver training, continuing education, and licensure related to these vehicles.

The development of driverless car technologies is critical for the advancement of the transportation industry. The majority of the promises and benefits will likely only be realized when all cars are equipped with these advanced technologies, enabling NHTSA’s Level 4 of fully autonomous driving. This is a tenuous period of transitioning new and unproved technologies into a complex sociotechnical system with significant variation in human ability.

In addition, public perception can become a major but surmountable obstacle. Great care should be taken, therefore, in experimenting with and implementing driverless technology—an ill-timed, serious accident could provoke unanticipated public backlash, which also could affect other robotic industries.

References