Spring 2022 www.uclajolt.com Vol. 27, No. 3

UCLA

JOURNAL OF LAW & TECHNOLOGY

AUTONOMOUS VEHICLE REGULATION & TRUST: THE IMPACT OF FAILURES TO COMPLY WITH STANDARDS

William H. Widen* & Philip Koopman**

ABSTRACT

The autonomous vehicle (AV) industry works very hard to create public trust in both AV technology and its developers. Building trust is part of a strategy to permit the industry itself to manage the testing and deployment of AV technology without regulatory interference. This article explains how industry actions to promote trust (both individually and collectively) have created concerns rather than comfort with this emerging technology. The article suggests how the industry might change its current approach to law and regulation from an adversarial posture to a more cooperative one in which a space is created for government regulation consistent with technology development. This article proposes a way forward that involves re-thinking the use of SAE J3016 as part of AV law and regulation, instead taking a new direction based on distinguishing test platforms from production vehicles.

* Professor, University of Miami School of Law, Coral Gables, Florida, currently researching laws and regulations relating to autonomous vehicles.

**Associate Professor of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania, specializing in autonomous vehicle safety.

TABLE OF CONTENTS

INTRODUCTION 171
I. TESLA'S FAILURE TO COMPLY WITH LAW 177
  A. *Importance of the SAE Level* 178
  B. *Why FSB Beta Vehicles are SAE Level 4* 180
    1. A Comparison of Tesla Statements with J3016 180
    2. Irrelevance of the Presence of a Human Driver to SAE Level ... 182
  C. *Tesla's Acknowledgment that Some of its Customers Beta Test* 184
  D. *Analysis of the Statute and Regulations* 184
  E. *Tesla's Communications with the DMV* 189
  F. *The Regulatory Corner* 190
II. AURORA'S FAILURE TO IDENTIFY A DEPLOYMENT STANDARD 192
  A. *Urgency of the AV Problem* 194
  B. *Absence of a Clear Deployment Standard Creates Uncertainty* 195
  C. *The Moral Hazard* 202
III. THE DIRTY DOZEN MYTHS ABOUT AV TECHNOLOGY 204
IV. "AUTONOMANDERING" AND THE CHALLENGE FOR THE POLITICAL PROCESS 216
V. PROBLEMS WITH J3016 AS A SAFETY STANDARD AND THE WAY FORWARD 223
  A. *Issues with J3016 to Define the Scope of a Law or Regulation* 223
    1. Regulatory Boundaries: Vehicles or Driving Automation Features? 224
    2. J3016 is Not a Safety Standard But "Something Else" 228
    3. The Apparent Failure of J3016 to Provide a Complete Vocabulary 232
  B. *An Illustration from the History of J3016* 235
  C. *The Way Forward* 241
    1. Differentiating Testing from Production 241
    2. Production Level 2 Regulation 242
    3. Test platform regulation involving safety drivers 248
  D. *An Alternative to SAE J3016* 251
CONCLUSIONS 253
  A. *Comparison to the Railroad Industry* 256
  B. *Recommendations to the AV Industry* 261

INTRODUCTION

The autonomous vehicle (AV) industry works very hard to create public trust in both AV technology and its developers.[1] Building trust is part of a strategy to permit the industry itself to manage the testing and deployment of self-driving vehicle technology without regulatory interference. The industry hopes that a high level of public trust will reduce or eliminate motivation for the federal government, states, and municipalities to enact meaningful safety laws and regulations governing both testing and deployment of AVs. The industry argues that increased regulation will impede innovation and slow the development of valuable technology.[2]

The AV industry supplements its appeals for trust by extolling a plethora of public benefits that AV technology supposedly will bring to society.[3] The AV industry attempts to shape both legislation and public opinion in various ways, including through Partners for Automated Vehicle Education (PAVE), a 501(c)(3) organization whose stated mission is to "educate" the public on the benefits of AV technology.[4]

In this article, we examine the behavior of certain AV industry participants in various concrete situations, considering whether actions by these AV industry participants promote trust or, conversely, create

1. *See, e.g.*, Aurora Innovation, Inc., Registration Statement (Form S-1), 83 (Nov. 5, 2021), https://www.sec.gov/ix?doc=/Archives/edgar/data/1828108/000119312521321663/d230050ds1.htm [hereinafter Aurora S-1] ("The opportunity to revolutionize transportation is massive, but this opportunity depends on trust. Our technology needs to be trustworthy. Our company needs to be trustworthy. And so our task is to build trust, one step at a time."); *see also* Reinvent Tech. Partners Y, Amendment No. 1 to Registration Statement (Form S-4), 244 (Aug. 27, 2021), https://www.sec.gov/ix?doc=/Archives/edgar/data/1828108/000119312521259448/d184562ds4a.htm [hereinafter Reinvent S-4] (filing for Aurora's predecessor).

2. *See, e.g.*, BILL CANIS, CONG. RSCH. SERV., R45985, ISSUES IN AUTONOMOUS VEHICLE TESTING AND DEPLOYMENT 8 (2021) [hereinafter ISSUES IN AV TESTING] (observing that "[p]roponents of autonomous vehicles note that lengthy revisions to current vehicle safety regulations could impede innovation, as the rules could be obsolete by the time they take effect."). We explain how regulation can be compatible with innovation despite industry protestations to the contrary. *See infra* text accompanying notes 132-137.

3. We discuss these advertised benefits in detail. *See infra* Section III (discussing myths about AVs).

4. *About*, PAVE, https://pavecampaign.org/about (last visited April 12, 2022).

concerns. To evaluate whether a particular action promotes trust, we consider both common sense intuitions as well as recommendations contained in three publications with intended global applicability: the recently published IEEE Standard Model Process for Addressing Ethical Concerns during System Design;[5] Ethics Guidelines for Trustworthy AI;[6] and Policy and Investment Recommendations for Trustworthy AI.[7] The European Commission set up the Independent High-Level Expert Group on Artificial Intelligence in June 2018 with the task of delivering these two reports on Trustworthy AI. While these European Commission reports primarily focus on development and use of AI in Europe, they intend a global reach.[8] These publications inform our analysis of actions and practices to identify those that merit trust and those that do not. (These publications are useful to inform policy and legislation even though none of them are in a form suitable for incorporation by reference into a law or regulation.)

Our analysis identifies certain trust destroying practices, suggesting that the public should not accept either AV industry appeals for trust or take at face value the proffered narrative of benefits. We illustrate our

---

5. IEEE SA, IEEE STANDARD MODEL PROCESS FOR ADDRESSING ETHICAL CONCERNS DURING SYSTEM DESIGN (2021) [hereinafter IEEE 7000], https://standards.ieee.org/ieee/7000/6781/ (available via purchase or subscription, on file with the authors). The Institute of Electrical and Electronics Engineers (or "IEEE") Standards Association has an express global reach with the stated mission of "Raising the World's Standards." See IEEE SA STANDARDS ASSOCIATION, https://standards.ieee.org/ (last visited April 12, 2022).

6. *High-Level Expert Group on Artificial Intelligence on Ethics Guidelines for Trustworthy AI* (Apr. 8, 2019), https://www.aepd.es/sites/default/files/2019-12/ai-ethics-guidelines.pdf [hereinafter *Trustworthy AI Guidelines*].

7. *High-Level Expert Group on Artificial Intelligence on Policy and Investment Recommendations for Trustworthy AI* (June 26, 2019), https://www.europarl.europa.eu/italy/resource/static/files/import/intelligenza_artificiale_30_aprile/ai-hleg_policy-and-investment-recommendations.pdf [hereinafter *Trustworthy AI Policy*].

8. "Beyond Europe, the Guidelines also aim to foster research, reflection and discussion on an ethical framework for AI systems at a global level." Trustworthy AI Guidelines, *supra* note 6, at 3. The Trustworthy AI Policy, *supra* note 7, is more Eurocentric than the Trustworthy AI Guidelines, though its section on law and regulation supplements the Trustworthy AI Guidelines, which focus on ethical principles and robustness. The discussion of legality is left for the Trustworthy AI Policy.

concerns by examining trust destroying practices of Tesla (who is not a member of PAVE), Aurora Innovation, Inc. (a member of PAVE who recently went public in a de-SPAC transaction),[9] and various narrative "myths" commonly used by AV industry members to advance their collective agenda. We continue by explaining how AV industry participants might destroy trust via political lobbying efforts undertaken to prevent municipalities from passing local safety laws and regulations by pre-emptive state legislation—and advise against it. AV industry participants have an active pre-emption effort underway in Pennsylvania and other states.

These failures at trust-building suggest that self-regulation by the AV industry is not a viable option and that more regulation is needed by federal, state, and local governments to promote safety. Despite our concerns, the AV industry might earn trust by changing its approach to law, regulation, and disclosure from an adversarial stance to a cooperative one, starting by compliance with standards promulgated by the engineering community itself—an action that, almost by definition, would not impede innovation.

Of course, the trust destroying practices we identify are merely illustrative. Nor is the industry's conduct the only threat to public confidence in AV technology: high-profile events—for example, Tesla's recent software rollouts and retractions of software for its Full Self-Driving (FSD) features,[10] the Uber accident in Arizona,[11] and the numerous videos circulating on the internet which show worrisome behavior by Tesla owners using Auto Pilot and FSD features[12]—can be equally corrosive to the public trust. With any new technology, accidents will occur despite the best efforts of industry and government to prevent them. The AV industry ought to take meaningful and visible steps to advance safety when it has

9. *See* Aurora S-1, *supra* note 1.

10. *See infra* notes 68–69 and accompanying text.

11. *See infra* note 155 and accompanying text.

12. *See, e.g.*, *infra* note 34 (Tesla refers to its current FSD deployments as "beta test" activities, however, that software is being used by non-employee drivers having no special tester training on public roads, and traffic laws still apply to such purported "testers").

the power to do so. It is important for the success of the AV industry that it adopt practices to build trust so that it can better thrive when adverse events beyond industry control materialize, as is inevitable.

As background, governmental approaches to AV regulation vary widely across the United States, with some jurisdictions (such as Arizona, Florida, Nevada, and Texas)[13] having enacted laws that give AV companies wide latitude to test and deploy AVs, whereas other states (such as California)[14] and municipalities (such as New York City)[15] seek to promote safety through law and regulation. Further, some states currently are considering laws and regulations governing AV testing and deployment.[16] The federal government has failed to enact any legislation, or promulgate regulations, governing AVs.[17]

This Article proceeds in five parts: Part I describes how Tesla fails to comply with California laws and regulations governing testing and deployment of AVs.[18] Both the Trustworthy AI Guidelines and the

---

13. Aurora S-1, *supra* note 1, at 88.

14. *See* CAL. VEH. CODE § 38750 (2017) [hereinafter CAL. CODE § 38750]; CAL. CODE REGS. tit. 13, §§ 227-228 (2008).

15. *See, e.g.*, N.Y.C., N.Y., RULES OF THE CITY OF NEW YORK tit. 34, ch. 4, § 4-17 (2021).

16. These include the states of Kansas (Kan. Reg. Leg. Sess. 2022, S.B. 379), Pennsylvania(Pa. Gen. Assem., Reg. Sess. 2021-2022, S. B. 965), and Washington (Wash. Reg. Leg. Sess., H.B. 1731). We understand preliminary discussions are underway to develop a joint effort at AV regulation among New England states. The approach taken in any future laws and regulations might be either permissive or protective. In some states, governors have issued executive orders relating to AVs. *See generally Autonomous Vehicles| Self-Driving Vehicles Enacted Legislation*, NAT'L CONF. OF STATE LEGISLATURES (Feb. 18, 2020), https://www.ncsl.org/research/transportation/autonomous-vehicles-self-driving-vehicles-enacted-legislation.aspx.

17. *See, e.g.*, ISSUES IN AV TESTING, *supra* note 2, at 1.

18. We introduced this idea in an essay for JURIST. *See* William H. Widen & Philip Koopman, *Do Tesla FSD Beta Releases Violate Public Road Testing Regulations?*, JURIST (Sept. 27, 2021, 12:00 PM), https://www.jurist.org/commentary/2021/09/william-widen-philip-koopman-autonomous-vehicles/. There is a growing awareness that Tesla's FSD beta releases may violate California law. *See* Bryant Walker Smith, *California's AV Testing Rules Apply to Tesla's "FSD"*, ROBOTICS.EE (Jan. 10, 2022), https://robotics.ee/2022/01/10/californias-av-testing-rules-apply-to-teslas-fsd/. Professor Bryant Walker Smith expressed similar concerns over Uber's testing in California in 2016. *See* Bryant Walker Smith, *Uber vs. The Law*, THE CTR. FOR INTERNET & SOC'Y AT STAN. L.

Trustworthy AI Policies identify compliance with law as one of three foundational elements needed to build trustworthy AI.[19]

Part II uses the SEC Registration Statement on Form S-1[20] for Aurora Innovation, Inc.'s (Aurora) public offering as an example of how the AV industry avoids committing to a specific safety performance standard prior to the initial deployment of AVs at scale.[21] Failure to disclose a standard for deployment violates principles of transparency and makes it impossible to effectively implement IEEE 7000.[22]

Part III identifies a "dirty dozen" myths about the status of AV technology, debunking key points used by the AV industry both to promote AV technology and argue for a light regulatory touch.[23] The use of these myths is inconsistent with concerns expressed in ethics standards that no one deceive the public about AI technology.[24]

SCH.: BLOG (Dec. 17, 2016, 7:47 am), https://cyberlaw.stanford.edu/blog/2016/12/uber-vs-law.

19. *See supra* note 6.

20. *See* Aurora S-1, *supra* note 1.

21. One of us previously suggested that this disclosure failure violates United States securities laws. *See* William H. Widen, *Autonomous Vehicles, Moral Hazards & the "AV Problem"* (Univ. of Mia. Legal Studies, Research Paper No. 3902217, 2021, forthcoming 3 Notre Dame J. Emerging Tech. Spring 2022), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3902217 (the focus in this article is on the impact of Aurora's disclosure choices on trust, not securities law compliance).

22. *See* IEEE 7000, *supra* note 5. Compliance with IEEE 7000 is voluntary. IEEE members, however, agree to comply with the IEEE Code of Ethics each year when renewing membership. *See infra* text accompanying note 121.

23. One of us previously identified various myths promoted by the AV industry as a technique to persuade elected officials to adopt a regulatory stance favorable to AV testing and deployment. *See* Philip Koopman, *Autonomous Vehicle Myths: The Dirty Dozen*, EETIMES (Oct. 22. 2021), https://www.eetimes.com/autonomous-vehicle-myths-the-dirty-dozen/#.

24. *See, e.g.,* IEEE CODE OF ETHICS, *infra* note 123 (noting with respect to technologies, including intelligent systems, there is an obligation "to be honest, and realistic in stating claims or estimates based on available data").

Part IV identifies a political problem we call "autonomandering"[25]—in which AV industry participants lobby elected officials in rural parts of a state (with less traffic density) to approve permissive and pre-emptive AV legislation which exposes constituents in urban areas (with greater traffic density) to a disproportionate increased risk of harm as compared with rural constituents. This problem provides an instance in which AV technology regulation presents a challenge for the democratic process—a general concern raised by the ethics standards that AI development and implementation not adversely impact either democracy or the rule of law.[26]

Part V presents an alternate path forward. It first identifies shortcomings in SAE J3016[27] as a safety standard. It then explains why a slightly modified version of SAE J3018[28] ought to be used (instead of J3016) as a foundation for law and regulation, with a focus on simplifying and clarifying the scope of laws and regulations governing AV safety in testing and deployment.

We conclude with an appeal to the AV industry for a shift in its approach to laws and regulation from an adversarial one to a cooperative one as the best method to promote and sustain valuable AV technology. One avenue for cooperation would be for the AV industry to engage in

---

25. We see this problem as related to the well-known problem of gerrymandering Congressional districts. Our terminology is inspired by Liza Dixon, *Autonowashing: The Greenwashing of Vehicle Automation*, TRANSP. RSCH. INTERDISC. PERSPS. (May 8, 2020), https://www.sciencedirect.com/science/article/pii/S2590198220300245

26. *See Trustworthy AI Guidelines*, *supra* note 6, at 11 (noting that "AI systems should serve to maintain and foster democratic processes and respect the plurality of values and life choices of individuals").

27. *See* SAE INT'L, TAXONOMY AND DEFINITIONS FOR TERMS RELATED TO DRIVING AUTOMATION SYSTEMS FOR ON-ROAD MOTOR VEHICLES J3016_202104 (2021) [hereinafter J3016:2021], https://www.sae.org/standards/content/j3016_202104/. We refer to prior versions of this publication as: J3016:2018, J3016:2016, and J3016:2014. We use J3016 without a following year of publication as a generic reference to the series of publications which all contain the same six level hierarchy of levels distinguishing different capabilities of driving automation systems or features. Reference to an "SAE Level" refers to a level described in J3016.

28. SAE INT'L, SAFETY-RELEVANT GUIDANCE FOR ON-ROAD TESTING OF PROTOTYPE AUTOMATED DRIVING SYSTEM (ADS)-OPERATED VEHICLES J3018_202012 (2020) [hereinafter J3018], https://www.sae.org/standards/content/j3018_202012/.

negotiated rulemaking with NHTSA to mandate AV industry compliance with applicable published engineering standards.[29]

I. TESLA'S FAILURE TO COMPLY WITH LAW

Tesla's behavior often fails to engender trust, most notably in its approach to testing its AV technology.[30] The AV industry generally conducts its public highway testing using specially trained employees as backup safety drivers. The outlier to this testing approach is the maverick, Tesla, who recently launched a wider distribution of its Full Self-Driving (FSD) suite of autonomy features for selected customers. Tesla's approach to safety eschews specially trained safety drivers, instead rolling out its testing product to a limited, but expanding, group of its customers who attain a sufficient "safe driver" score on a metric internally created by Tesla.[31] For some, the Tesla approach shows a similar concern with safety, despite the absence of trained safety drivers.

For others, Tesla's approach is disturbing. Tesla's testing approach is problematic for many reasons, but the original sin relates to what we perceive as a deliberate misapplication of California law and regulations by mischaracterizing its FSD beta features as SAE Level 2 when they really qualify as more heavily regulated Level 4 under J3016. In fact, Tesla fails to comply with law because the law and regulations, if read properly, do not allow permit-less testing of autonomous vehicles with untrained drivers.

---

29. *See, e.g.*, MAEVE P. CAREY, CONG. RSCH. SERV., R46756, NEGOTIATED RULEMAKING: IN BRIEF 8 (2021) [hereinafter NEGOTIATED RULEMAKING]. Indeed, NHTSA has itself recently suggested the "new approach[]" of using negotiated rulemaking in a partnership with industry. NAT'L HIGHWAY TRAFFIC SAFETY ADMIN. [NHTSA], NHTSA 2020 REPORT: PEOPLE SAVING PEOPLE (2020) [hereinafter NHTSA 2020 REPORT], https://one.nhtsa.gov/nhtsa/whatis/planning/2020Report/2020report.html.

30. Tesla has taken other actions which hinder development of trust such as disseminating safety statistics that did not hold up under scrutiny. *See, e.g.*, Edward Niedermeyer, *NHTSA's Flawed Autopilot Safety Study Unmasked*, THE DRIVE (Feb. 11, 2019), https://www.thedrive.com/tech/26455/nhtsas-flawed-autopilot-safety-study-unmasked (noting that "Tesla repeatedly puts out easily-debunked statistics and conceals its data in a system with as little transparency and accountability as possible").

31. *See Support: Safety ScoreBeta*, TESLA, https://www.tesla.com/support/safety-score (last visited April 12, 2022).

Both the Trustworthy AI Guidelines and the Trustworthy AI Policies identify compliance with law as one of three foundational elements needed to build trustworthy AI.[32] If this misapplication of law and regulation remains unchallenged, the risk remains that other AV industry participants, not only Tesla, may use this "loophole" to gain some advantage at the expense of safety[33] (though we do not foresee other major AV industry participants going so far as to use their own customers as "beta testers").

A. Importance of the SAE Level

One argument in support of legal compliance by Tesla's FSD beta vehicles relies on classification of the FSD beta features as SAE Level 2. On this reasoning, AVs must, by definition, qualify as SAE Level 3, 4 or 5; and, only AVs (as so defined) are subject to these laws. Thus, by maintaining an SAE Level 2 classification, Tesla hopes that FSD beta will avoid meaningful regulation. When convenient, Tesla promotes the view that its vehicles' features, including FSD beta, only qualify for SAE Level 2 classification.

In correspondence, Tesla has suggested this classification to the California Department of Motor Vehicles (DMV) for its self-driving technologies.[34] Publicly available testing videos for FSD beta vehicles

---

32. *See Trustworthy AI Guidelines*, *supra* note 6, at 6; *Trustworthy AI Policy*, *supra* note 7, at 37. In those jurisdictions in which a rolling stop violates traffic law, Tesla's Full Self-Driving (FSD) beta may provide another instance of failure to comply with law. *See* Emma Roth, *Tesla's 'Full Self-Driving' Beta Has an 'Assertive' Driving Mode That 'May Perform Rolling Stops'*, THE VERGE (Jan. 9, 2022, 7:12 PM), https://www.theverge.com/2022/1/9/22875382/tesla-full-self-driving-beta-assertive-profile; Kea Wilson, *Why Tesla Can Program Its Cars to Break Road Safety Laws*, STREETSBLOG USA (Jan. 12, 2022), https://usa.streetsblog.org/2022/01/12/why-tesla-can-program-its-cars-to-break-road-safety-laws/.

33. Another AV company, Embark Technology, Inc., may have violated California law and regulations when it drove a truck autonomously on a circuit of roads around Oakland as part of a due diligence demonstration for a public offering. *See* Northern Genesis Acquisition Corp. II, Amendment No. 1 to Registration Statement (Form S-4), 78-79 (Aug. 30, 2021) [hereinafter Embark S-4] (disclosing a technology demonstration by Embark for its de-SPAC transaction), https://www.sec.gov/ix?doc=/Archives/edgar/data/1827980/000110465921111215/ngab-20210830xs4a.htm.

34. *See California DMV Tesla Robo-Taxi / FSD E-Mails*, PLAINSITE (Mar. 5, 2021), https://www.plainsite.org/documents/242a2g/california-dmv-tesla-robotaxi--fsd-emails/ (posting a response to a public records request).

suggest,[35] however, that these beta test drivers operate their vehicles *as if* to validate SAE Level 4 features, often revealing dramatically risky situations created by use of the vehicles in this manner. CNN recently independently confirmed that operation of FSD technology represents a hazard based on its own use of Tesla vehicles in Brooklyn.[36] Lawmakers and regulators should focus on this reality and recognize that FSD beta testing constitutes SAE Level 4 testing on public roads. Because of this reality, FSD beta testers should be subject to the same regulatory oversight as all other Level 4 testers to ensure the safety of road users and bystanders.

Moreover, seen in its true light, the sale of FSD beta vehicles may constitute an unlawful deployment of AVs without applicable safety standard compliance in California and, perhaps, some other jurisdictions. (An alternate argument could be made that FSD beta testing is effectively SAE Level 3 operation since that involves a subset of Level 4 capabilities, but the net effect is the same.)

A great many aspects of regulation depend on SAE level. Only "autonomous vehicles" are subject to specific statutory requirements on the operation and deployment of autonomy features in California.[37] A vehicle does not qualify as an "autonomous vehicle" merely because it has driver

---

35. *See, e.g.*, Jake Lingeman, *Tesla's 'Full-Self Driving' Update 10 Is Still Pretty Scary*, CARBUZZ (Sept. 14, 2021), https://carbuzz.com/news/teslas-full-self-driving-update-10-is-still-pretty-scary. There have been repeated incidents of Tesla FSD beta vehicles attempting to turn left into oncoming traffic at significant speed (above 10 mph). For example, defective turning behavior reproduced in both FSD beta 10.3 and FSD beta 10.4 with the same driver, the same, vehicle, and the same left turn. *See* Phil Koopman (@PhilKoopman), TWITTER (Nov. 9, 2021, 7:25 AM), https://twitter.com/PhilKoopman/status/1458063125194936320. In this case a left turn was intended but was commanded by the automation despite detected oncoming traffic. Also, FSD exhibited defective turning behavior that deviates from a straight trip route. In this case the FSD beta 10.4 system fails to detect an illuminated "no left turn" traffic sign. *See* Phil Koopman (@PhilKoopman), TWITTER (Nov. 17, 2021, 7:43 AM), https://twitter.com/PhilKoopman/status/1460966916617641987.

36. *CNN Tried Tesla's 'Full Self-Driving' Mode on NYC Streets. It Didn't Go Great*, CNN BUS. (Nov. 17, 2021), https://www.cnn.com/videos/business/2021/11/17/tesla-3-full-self-driving-mode-test-mc-zw-orig.cnn-business (posting video of CNN team using FSD on a busy street in Brooklyn).

37. *See* CAL. CODE § 38750, *supra* note 14.

assistance features, such as collision avoidance systems.[38] The current California statute and DMV regulations specifically reference the SAE taxonomy for driving automation systems, limiting the scope of the term 'autonomous vehicle' to Levels 3, 4 and 5.[39]

B. *Why FSB Beta Vehicles are SAE Level 4*

1. A Comparison of Tesla Statements with J3016

A comparison of Tesla's public statements with SAE standards document J3016 establishing the criteria for assigning a level to an automated vehicle demonstrates that FSD beta testing constitutes SAE Level 4 testing on public roads. Consider first Tesla's description of its Full Self Driving Capability:

> All new Tesla cars have the hardware needed in the future for full self-driving in almost all circumstances. *The system is designed to be able to conduct short and long distance trips with no action required by the person in the driver's seat*.
>
> The future use of these features without supervision is dependent on achieving reliability far in excess of human drivers as demonstrated by billions of miles of experience, as well as regulatory approval, which may take longer in some jurisdictions. As these self-driving capabilities are introduced, your car will be continuously upgraded through over-the-air software updates.[40]

SAE J3016:2021[41] defines Level 4 capability as "[t]he sustained and ODD-specific performance by an ADS of the entire DDT and DDT fallback

38. *See* CAL. CODE § 38750(a)(2)(B), *supra* note 14.

39. CAL. CODE § 38750, *supra* note 14; CAL. CODE REGS., *supra* note 14. The California regulation incorporates the 2016 version of the SAE taxonomy by reference whereas the newly amended § 38750 incorporates the 2021 version. By its terms, J3016:2021, *supra* note 27, supersedes prior versions of the taxonomy, which has remained essentially the same across versions, with levels of autonomy capability from Level 0 to Level 5. The important concept of "design intent" was introduced in J3016:2016, the first revision, and continues in the 2018 and 2021 revisions. *See infra* text accompanying notes 198–99.

40. *Future of Driving*, TESLA (emphasis added), https://www.tesla.com/autopilot (last visited April 12, 2022).

41. We refer to J3016:2021, *supra* note 27, in our discussion. On our analysis, this most recent version is substantively equivalent in all relevant aspects to J3016:2016 used by the

without any expectation that a user will need to intervene."[42] Further, "[t]he level of a driving automation system feature corresponds to the feature's production design intent."[43]

As shown by the following table, Tesla's description of its FSD capability matches the SAE J3016:2021 requirements for Level 4. An explanation of SAE J3016:2021 terms follows the table.

SAE J3016 Requirement	Tesla Description
"sustained"	"conduct short and long distance trips"
"ODD-specific performance"	"almost all circumstances"
"by an ADS"	"All new Tesla cars have the hardware needed" and "software updates"
"the entire DDT"	"conduct short and long distance trips with no action required by the person in the driver's seat"
"DDT Fallback"	"conduct short and long distance trips with no action required by the person in the driver's seat"
"without any expectation that a user will need to intervene"	"no action required by the person in the driver's seat"
"design intent"	"The system is designed to be able to conduct"

The requirement that performance must be "sustained" is distinguished from momentary intervention during potentially hazardous situations, such as electronic stability control and automated emergency braking, and certain types of driver assistance systems, such as lane keeping assistance,

DMV in the California Code of Regulations. *See infra* notes 187-199 and accompanying text.

42. J3016:2021, *supra* note 27, at 17 tbl.1. The definition in J3016:2016 is identical, but we refer to the newest released version of the standard in this discussion because the California legislature has recently incorporated it by reference in its autonomous vehicle statute, even though the DOT regulations still refer to J3016:2016. Relevant differences between the 2016 and 2021 versions are discussed later in Part V, but do not change our analysis.

43. J3016:2021, *supra* note 27, § 8.2, at 36.

because these features do not perform part or all the Dynamic Driving Task (DDT) on a sustained basis.

"ODD" stands for "Operational Design Domain" which is the environment and other limited circumstances in which a Level 4 vehicle is intended to operate. By way of contrast, a Level 5 vehicle is designed to operate in all circumstances.

"ADS" stands for "Automated Driving System". An ADS performs the automated driving task, comprising both computer hardware and software. By way of contrast, a driver assistance feature, such as cruise control, does not drive the vehicle but merely assists the driver. Tesla's description of its FSD capability initially omits a reference to software, perhaps attempting to distinguish its product from Level 4. However, hardware alone does not comprise an ADS and the later reference to "software updates" confirms the presence of initial software and an intent to continuously upgrade FSD capabilities.

"DDT" stands for "Dynamic Driving Task". The DDT includes steering and speed control, but not destination selection. To perform this task, the ADS which supports the DDT must, among other things, monitor the driving environment by object and event detection, recognition, and response formulation.

"DDT Fallback" stands for the process of bringing a vehicle to a safe state (for example, stopping on the shoulder of a road) following a failure of some aspect of the ADS, as well as the occurrence of other conditions reasonably expected for some trips (for example, a broken axle). In a Level 4 vehicle, the DDT Fallback is handled by the vehicle, not a human driver. Even if Tesla's eventual deployment contemplates that its users are expected to handle DDT Fallback, at most that merely reduces the FSD beta vehicle to Level 3—a level which is still subject to regulation as an AV.

Based on this analysis, Tesla's description of the FSD's intended design capability clearly describes an SAE Level 4 feature. Tesla ought not avoid regulation by the label it self-assigns to its vehicles.

2. Irrelevance of the Presence of a Human Driver to SAE Level

The presence of a human driver does not preclude an FSD beta vehicle from Level 4 classification, as the current version of J3016 makes clear in

Section 8.2 (a point emphasized by one of the authors elsewhere as "Myth 10" about using the SAE Levels to classify vehicle automation)[44]:

> The level of a *driving automation system* feature corresponds to the *feature's* production design intent. This applies regardless of whether the *vehicle* on which it is equipped is a production *vehicle* already deployed in commerce, or a test *vehicle* that has yet to be deployed. As such, it is incorrect to classify a Level 4 design-intended *ADS feature* equipped on a test *vehicle* as Level 2 simply because on-road testing requires a test *driver* to supervise the *feature* while engaged, and to intervene if necessary to maintain *operation*.[45]

The SAE J3016:2021 Section 8.2 criteria for assigning SAE Level 4 hinge on design intent. If the manufacturer's design intent is Level 4, then it is a Level 4 vehicle even if there is a test driver to supervise while the feature is engaged and intervene when necessary. Significantly, a vehicle can qualify as Level 4 even if its hardware and software are not a particularly competent or safe instantiation of Level 4 technology.[46] That, we suggest, is the reality of the current situation, and why regulatory oversight of FSD beta is critical.

Tesla's description of the FSD feature makes it quite clear that Tesla has Level 4 design intent: "The system is designed to be able to conduct short and long distance trips with no action required by the person in the driver's seat."[47] In contrast, at Level 2 the driver is required to "complete the OEDR subtask" portion of the DDT, which involves Object and Event Detection and Response.[48] The whole point of FSD, as generally represented by Tesla's marketing materials and public messaging, is that the driver no longer

---

44. Philip Koopman, *SAE J3016 User Guide*, CARNEGIE MELLON UNIV., https://users.ece.cmu.edu/~koopman/j3016/#myth10 (last updated Sept. 4, 2021); *see also infra* Part III.

45. J3016:2021, *supra* note 27, § 8.2, at 36.

46. A safe, competent Level 4 design is one where the driver should be able to literally go to sleep during the journey and expect to be acceptably safe even if equipment fails, and not be under any burden to monitor or take over operation to ensure safety. By way of contrast a vehicle that requires frequent driver intervention to avoid a fatal crash is a Level 4 vehicle so long as the design intent is to eventually get better.

47. *See Future of Driving*, *supra* note 40.

48. J3016:2021, *supra* note 27, at 17 tbl. 1.

needs to drive (that is, the FSD feature actually fully self-drives), which necessarily removes the OEDR subtask burden from the human driver.

C. Tesla's Acknowledgment that Some of its Customers Beta Test

As to whether the person in the driver's seat qualifies as a "test driver," Tesla itself is calling such drivers FSD "beta testers." Tesla has been accepting and granting electronic applications for testers via a beta test request button, and has been giving access selectively, making such further distribution an expansion of a test program rather than a wide public release.

While Tesla hopes to reassure the public by saying that only good drivers will receive permission to test FSD beta, this only reinforces the notion that FSD beta is a selectively released pre-production test system, and not a road-ready full production feature. In other words, Tesla is having selected but untrained civilian drivers do on-road testing of their "beta" SAE Level 4 FSD features. This combination of vehicle plus amateur test driver arrangement has been documented to drive recklessly and otherwise behaves dangerously on public roads.[49]

When the FSD beta vehicle is properly recognized as a Level 4 capable vehicle, testing becomes a problem under the California statutes and regulations because this beta testing does not comply with law, as outlined in the next subsection.

D. Analysis of the Statute and Regulations

Under California law, "'[a]utonomous vehicle' means any vehicle equipped with autonomous technology that has been integrated into that vehicle that meets the definition of Level 3, Level 4, or Level 5 of SAE International's 'Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles, standard J3016 (APR2021),' as may be revised."[50] And "'[a]utonomous technology' means

---

49. Even worse, there is no enforcement mechanism to ensure that the driver who "earned" the safety score is actually behind the wheel during "testing." As an example, one social media video states that a driver borrowed his neighbor's Tesla with FSD beta and shows him running a stop sign without even slowing down. *See* Kyle Conner (@itskyleconner), TWITTER (Oct. 26, 2021, 1:41 PM), https://twitter.com/itskyleconner/status/1453069194799501323.

50. 2021 Cal. Legis. Serv. Ch. 277 § 1(a)(2)(A) (S.B. 500) (West) (updating CAL. CODE § 38750, *supra* note 14). Before the September 23 amendment, the term 'Autonomous

technology that has the capability to drive a vehicle without the active physical control or monitoring by a human operator."[51] A Tesla vehicle with FSD beta satisfies these definitions because, based on Tesla's marketing statements—and various videos posted online by Tesla FSD beta testers[52]—FSD-beta-equipped vehicles are capable of driving without active physical control or monitoring. Though Tesla's instructions stipulate that the human driver must constantly monitor driving, the instruction does not make FSD beta vehicles any less capable of driving without human control or monitoring.[53]

The law is about "capability" and not about the instructions in a manual. FSD beta goes beyond the sort of collision avoidance or driver assistance systems that do not make a vehicle "autonomous."[54]

Vehicle' meant "any vehicle equipped with autonomous technology that has been integrated into that vehicle." *See* CAL. CODE § 38750, *supra* note 14.

51. CAL. CODE § 38750(a)(1), *supra* note 14.

52. Mahmood Hikmet, *Tesla FSD Beta Danger Compilation*, YOUTUBE (Nov. 11, 2021), https://www.youtube.com/watch?v=GmoroFK1A_o.

53. *See, e.g.*, *Support: Full Self-Driving Capability Subscriptions*, TESLA, https://www.tesla.com/support/full-self-driving-subscriptions#faq (last visited April 12, 2022) ("Note: These features are designed to become more capable over time; however the currently enabled features do not make the vehicle autonomous. The currently enabled features require a fully attentive driver, who has their hands on the wheel and is prepared to take over at any moment.").

54. *See* CAL. CODE § 38750(a)(2)(B), *supra* note 14 ("An autonomous vehicle does not include a vehicle that is equipped with one or more collision avoidance systems, including, but not limited to, electronic blind spot assistance, automated emergency braking systems, park assist, adaptive cruise control, lane keep assist, lane departure warning, traffic jam and queuing assist, or other similar systems that enhance safety or provide driver assistance, *but are not capable, collectively or singularly, of driving the vehicle without the active control or monitoring of a human operator.*" (Emphasis added)) The exception appears simply to make clear that traditional driver assistance systems do not render a vehicle "autonomous" because they do not drive, but merely assist. Again, the key is capability, and FSD beta has this capability.

So, by statute, the Tesla FSD beta is an "autonomous vehicle" because of its capabilities which satisfy the SAE criteria for Level 3 or 4. The California DMV regulations contain a further clarification for an "autonomous *test* vehicle":

> For the purposes of this article, an "autonomous test vehicle" is equipped with technology that makes it capable of operation that meets the definition of Levels 3, 4, or 5 of the SAE International's *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles, standard J3016 (SEP2016)*, which is hereby incorporated by reference.[55]

Thus, if Tesla FSD beta is merely Level 2, it is not an autonomous test vehicle and would not be an autonomous vehicle subject to regulation. But the picture changes when FSD beta vehicles are properly recognized as Level 3 or 4 because its "beta testing" program no longer complies with law.

A statutory autonomous vehicle may be operated on public roads in California for "testing purposes" by a driver possessing the proper class of license if three conditions are met:

> 1. The autonomous vehicle is being operated on roads in this state solely by employees, contractors, *or other persons designated by the manufacturer* of the autonomous technology.
>
> 2. The driver shall be seated in the driver's seat, monitoring the safe operation of the autonomous vehicle, and capable of taking over immediate manual control of the autonomous vehicle in the event of an autonomous technology failure or other emergency.
>
> 3. Prior to the start of testing in this state, the manufacturer performing the testing shall obtain an instrument of insurance, surety bond, or proof of self-insurance in the amount of five million dollars ($5,000,000), and shall provide

---

55. CAL. CODE REGS. tit. 13, § 227.02(a)(2) (2008); *accord* CAL. CODE REGS. tit. 13, § 228.02 (b) (2008) (stating that the definition of "autonomous vehicle" meets SAE Levels 3, 4, or 5). An argument could be made that every Tesla vehicle with the hardware required to support FSD, including especially every vehicle for which customers have paid for FSD even if not yet authorized for beta operation, meets this requirement because it is equipped with the hardware technology to provide that capability even if not enabled by software.

> evidence of the insurance, surety bond, or self-insurance to the department in the form and manner required by the department pursuant to the regulations adopted pursuant to subdivision (d).[56]

Tesla has satisfied subsection (3) by virtue of having obtained a license in California to test with a driver.[57] Tesla will argue that it satisfies subsection (2) by virtue of certain of Tesla's statements, including in owner's manual instructions to its FSD beta customers to stay alert, ready to always take over control of the vehicle. And with respect to subsection (1), Tesla will argue that it has designated its customers to do the testing because of its selective rollout of FSD beta and its qualifications which allow only a limited number of its customers to participate in the "beta testing."

But it is not clear that Tesla's FSD beta customers do, in fact, qualify as "designees": Pursuant to regulation, "'[d]esignee' means the natural person *identified by the manufacturer to the department* as an autonomous vehicle test driver authorized by the manufacturer to drive or operate *the manufacturer's autonomous test vehicles* on public roads."[58]

Thus, there are two considerations. First, has Tesla sufficiently identified its FSD beta customers to the DMV as "autonomous vehicle test drivers"? We are not aware of Tesla making any such designation of test drivers to the DMV (unless Tesla's public remarks qualify). Second—and more importantly—are Tesla's FSD beta testers operating the "manufacturer's autonomous test vehicles?" It seems that they are not: the FSD beta customers are operating their own vehicles, and not those that belong to Tesla. The regulatory scheme contemplates that testing of autonomous technology be limited to a manufacturer's vehicles, and that such technology would be deployed to the public only after testing had been completed. Perhaps the regulatory scheme looks the way it does because it never occurred to the legislature or the DMV that any manufacturer would be so bold (or reckless) as to use its own customers as test drivers.

---

56. CAL. CODE § 38750(b), *supra* note 14 (emphasis added).

57. *Autonomous Vehicle Testing Permit Holders*, STATE OF CAL. DMV, https://www.dmv.ca.gov/portal/vehicle-industry-services/autonomous-vehicles/autonomous-vehicle-testing-permit-holders/ (last visited April 12, 2022).

58. CAL. CODE REGS. tit. 13, § 227.02(e) (2008) (emphasis added).

Further, California regulations provide minimum qualifications for autonomous vehicle test drivers, including three years of licensure, not more than one violation point count, not having been at fault in any accident resulting in injury or death, no convictions in the prior 10 years for driving under the influence of alcohol or drugs, and completion of the manufacturer's autonomous vehicle test driver program.[59] Because Tesla does not verify the driving records of its FSD beta testers, and does not require them to complete an autonomous vehicle test driver program, it seems unlikely that any FSD beta tester would qualify as a designee.

To the extent that Tesla's FSD beta customers do not qualify as designees, they cannot be involved in "testing" FSD features; that is, they cannot operate "an autonomous vehicle on public roads . . . for the purpose of assessing, demonstrating, and validating the autonomous technology's capabilities."[60] If an autonomous vehicle is not being operated for testing purposes, it "shall not be operated on public roads until the manufacturer submits an application" to the DMV, and the DMV approves it.[61] We are not aware of any such application or approval. Thus, it seems that, with respect to its FSD beta testing program, Tesla is deploying autonomous vehicles in violation of regulations.[62]

California's statutory and regulatory schemes appear designed to facilitate the development of new technologies while protecting the public from the dangers posed by immature technologies. In using its customers to test FSD beta, Tesla disregards these policy judgments. Pursuant to regulation, an autonomous vehicle shall not be deployed on any public road in California until the manufacturer has submitted, and the DMV

---

59. *Id.* § 227.34.

60. *Id.* § 227.02(o).

61. CAL. CODE § 38750(c), *supra* note 14. The required contents of an application are set forth by statute. *See id.*

62. CAL. CODE REGS. tit. 13, § 228.02(c) (2008) (defining "deployment" of an autonomous vehicle as "the operation of an autonomous vehicle on public roads by members of the public who are not employees, contractors, or designees of a manufacturer or for purposes of sale, lease, providing transportation services or transporting property for a fee, or otherwise making commercially available outside of a testing program.").

approved, an Application for a Permit to Deploy Autonomous Vehicles on Public Streets, form OL 321 (Rev. 7/2020).[63]

E. Tesla's Communications with the DMV

Our conclusion does not change based on representations that Tesla made to California regulators to the effect that its vehicles are SAE Level 2, a classification presumably made to avoid regulatory oversight and permitting processes required of more highly automated vehicles, including Level 4 vehicles.

An analysis of released e-mails between Tesla and the California DMV reveals that Tesla left itself room to maneuver by careful word choice.[64]

- Tesla promised "we won't deploy any autonomous vehicle feature without a deployment permit."[65] However, Tesla might not consider a "test" program to be a "deployment", so this statement does not necessarily apply to FSD beta.
- FSD is a distinct feature from AutoPilot (AP). AP is included standard in all newer Tesla vehicles, whereas FSD requires an additional fee, confirming that there are two separate products. Thus, statements regarding AP being Level 2 do not necessarily bear on FSD because they are distinct and different product features. (SAE J3016:2021 states that a Level is associated with a feature, not the entire vehicle. AP can be at Level 2 while FSD is at Level 4.)
- The Tesla letter of November 20, 2020,[66] limits its discussion to current capabilities, and not design intent, whereas design intent is the crux of SAE levels. (It is worth noting that the letter refers to "the small handful of non-employee drivers in the pilot."[67] This number

---

63. *Id.* § 228.06(a) (governing post-testing deployment).

64. *See* PLAINSITE, *supra* note 34.

65. Email from Al Prescott, Chief Legal Officer, Tesla, to Brian G. Soublet, Deputy Dir. & Chief Couns., Cal. DMV (Dec. 20, 2019 11:17 AM) (available at PLAINSITE, *supra* note 34).

66. Letter from Eric C. Williams, Assoc. Gen. Couns. - Reg., Tesla, to Miguel Acosta, Chief, Autonomous Vehicles Branch - Cal. DMV (Nov. 20, 2020) (available at PLAINSITE, *supra* note 34).

67. *Id.* at 3; *see also* Elon Musk (@elonmusk), TWITTER (Sept. 17, 2021, 8:43 PM), https://twitter.com/elonmusk/status/1439042334155497474 (claiming that 2000 beta users had been operating in the year following that letter.

increased to almost 12,000 beta testers as of October 29, 2021,[68] and had increased to approximately 60,000 as reflected in Tesla's fourth quarter and fiscal year 2021 update, with potentially many more coming soon.[69])

- The closest Tesla comes to an SAE Level statement is the imprecise notion that Tesla "continues to firmly root the vehicle in SAE Level 2 capability."[70] But that is not a statement that the technology is Level 2. It means that Tesla's path to Level 4 starts at Level 2. That simply reflects the reality of an evolution in capabilities from AP (which is Level 2) to FSD (which is really Level 4).
- The Tesla letter of December 14, 2020[71] refers to a "final release" and release "to the general public" being SAE Level 2, rather than characterizing the level of *current beta releases* to selected testers. Indeed, Tesla might never issue a "final release," instead keeping FSD in beta indefinitely, offering the feature to essentially all "qualified" Tesla owners, thus technically avoiding a "deployment." Or it might rebrand FSD one day and declare the functionality formerly known as FSD to then be a "new" Level 4 feature.

A complete analysis of the disclosed documents posted at PLAINSITE is beyond the scope of this Article. But we were unable to find any unambiguous statement by Tesla that the FSD beta program is at SAE Level 2, as opposed to the characterization of the anticipated "final release." (In any event, any such statement about FSD beta, if made, would be incorrect on our analysis.)

F. The Regulatory Corner

Tesla has painted itself into a regulatory corner. If Tesla denies a design intent that its FSD beta feature satisfies SAE Level 4 capability, Tesla's

---

68. NHTSA, PART 573 SAFETY RECALL REPORT, 21V-846, OMB CONTROL NO.: 2127-0004 (2021) (noting 11,704 vehicles potentially affected), https://static.nhtsa.gov/odi/rcl/2021/RCLRPT-21V846-7836.PDF.

69. TESLA, Q4 AND FY2021 UPDATE at 10, available at https://tesla-cdn.thron.com/static/WIIG2L_TSLA_Q4_2021_Update_O7MYNE.pdf. This number will increase as Tesla grants more customers access to FSD technology.

70. *See* PLAINSITE, *supra* note 34; *see also* Letter from Eric C. Williams to Miguel Acosta, *supra* note 66. In our view, Tesla's wording is deliberately vague by suggesting that vehicles are Level 2 without making an express statement to this effect. The express statement might be false, as we suggest in the case of FSD.

71. *See* Letter from Eric C. Williams to Miguel Acosta, *supra* note 66.

pervasive statements and messaging strategy to customers purchasing FSD—suggesting that FSD-equipped vehicles are, in fact, capable of full self-driving-—would be misleading. Tesla simply must have the design intent to develop and perfect Level 4 technology (and be in the process of honoring its promises to its customers by testing Level 4 features). To produce truly safe Level 4 technology for general release to the public, common practice would first test Level 4 technology that is less capable. Indeed, Level 4 performance would be expected to improve over time. For regulatory purposes, given the applicable statutory definitions, it simply will not suffice to deny actual design intent to build either a Level 3 or Level 4 vehicle because those definitions turn, in the first instance, on capability—which the FSD beta vehicle possesses at Level 4 and Tesla must test prior to a full public deployment.

The only thing that saves Tesla from the California scheme of regulatory oversight is the willingness of the California regulators, for whatever reasons, to continue to take Tesla's classification of its FSD technology as Level 2 at face value.[72] But as explained above, there is every reason to reject such a classification by carefully parsing the language of J3016, together with the statutory and regulatory definitions.

For the reasons outlined above, state departments of transportation around the United States should classify the Full Self-Driving beta releases as an SAE Level 4 feature, with applicable regulatory and operational guidance applied accordingly based on individual state laws and regulations.[73]

---

72. *See* Hyunjoo Jin, *San Francisco Raises Tesla 'Self-Driving' Safety Concerns As Public Test Nears*, REUTERS (Sept. 23, 2021, 5:57 PM), https://www.reuters.com/business/autos-transportation/san-francisco-raises-tesla-self-driving-safety-concerns-public-test-nears-2021-09-23/ (noting that California's state regulator recently said: "Based on information Tesla has provided the DMV, the feature does not make the vehicle an autonomous vehicle per California regulations."). Upon the request of California state Senator Gonzalez, the California DMV has announced that it is re-evaluating its positions with respect to Tesla. *See* Letter from Steve Gordon, Dir., Cal. DMV, to Hon. Lena A. Gonzalez, Chair, Senate Comm. on Transp. (Jan. 7, 2022) (available at Andrew J. Hawkins, *California Is 'Revisiting' Tesla's Full Safe-Driving Beta In Light of 'Dangerous' Videos*, THE VERGE (Jan. 12, 2022, 1:42 PM), https://www.theverge.com/2022/1/12/22880305/tesla-fsd-beta-california-dmv-autonomous-vehicle-letter).

73. Though we focus on laws and regulations in California in this article, other states and the District of Columbia have adopted laws and regulations which are capable of

Testing potentially dangerous products on public highways cannot, as a matter of policy, properly be addressed as an exercise in wordplay, labeling sophisticated technology as merely Level 2 while looking the other way. Public safety—and the earning of public trust—requires more. The fact that Tesla approaches safety regulation as a classification game to be won or lost, without considering the safety consequences of winning this game, provides a reason to withhold trust. Flouting the application of duly enacted laws and regulations provides a signature example of a trust destroying failure to comply with law.

II. AURORA'S FAILURE TO IDENTIFY A DEPLOYMENT STANDARD

The AV industry must answer a practical ethical question: How will a company know when its AV technology is safe enough to deploy at scale?[74] We call this question, the "AV Problem." Aurora Innovation, Inc. (Aurora) and the AV industry more generally, might use IEEE 7000[75] to address this problem. Nevertheless, as explained in this Part II, neither Aurora nor other AV industry participants wish to publicly state how they will address the AV Problem. What level of safety does the AV industry aim to achieve for a first deployment?

This Part considers the AV Problem through the lens of a November 5, 2021, filing by Aurora of a registration statement on Form S-1[76] with the Securities and Exchange Commission. Aurora hopes to be a leader in systems for AVs.[77]

The Aurora S-1 reveals a potentially significant material omission: it fails clearly to disclose Aurora's internal standard for initially deploying

---

manipulation by using what we call the "Level 2 loophole" because they define an automated vehicle by reference to SAE Levels 3, 4 or 5. *See, e.g.*, D.C. CODE ANN. § 50-2351(1A) (West 2021).

74. *See, e.g.*, Patrick McGee, *Robotaxis: Have Google and Amazon Backed the Wrong Technology?*, FIN. TIMES (July 18, 2021), https://www.ft.com/content/46ff4fe4-0ae6-4f68-902c-3fd14d294d72 (subscription required) ("Since Google launched its self-driving car project in 2009, the biggest challenge has been one of technology: can it be safe enough to deploy at scale?").

75. *See* IEEE 7000, *supra* note 5.

76. *See* Aurora S-1, *supra* note 1.

77. *Id.* at 85 (describing the plan to be a global leader in self-driving technology); *see also* Reinvent S-4, *supra* note 1, at 245.

AVs at scale. Development of technology satisfying a more stringent safety standard takes longer to develop than technology meeting a lesser standard. The Aurora S-1 makes clear that Aurora must deploy AV technology quickly for financial success. For this reason, Aurora's deployment standard is material and its omission a potential violation of securities laws.[78] Beyond the apparent securities law violation (another trust destroying feature), however, the failure to identify the applicable standard for deployment makes it impossible for Aurora to satisfy the requirements of IEEE 7000.

IEEE 7000 aims to support companies in creating ethical value through system design. "Creating ethical value is a vision for organizations that recognizes their central role in society as shapers of well-being and carriers of societal progress that benefits humanity. Implementing IEEE Std 7000 can help [a company] to strengthen [its] value proposition and avoid value harm."[79] IEEE 7000 supports an organization's efforts to behave ethically and create ethical value through system design by setting forth internal processes and procedures conducive to production of ethical results and promoting ethical treatment of persons. This allows technologists to "align products and services with the results valued by acquirers, consumers, and users."[80]

IEEE 7000 applies to all kinds of products and services, including AI systems—the category into which AV technology falls. It envisions a "Case for Ethics"[81]—which is like a safety case[82]—to provide a structured account of the ethical and technical activities undertaken while pursuing an ethically aligned design for a system of interest. It serves as a project memory and an auditable repository. It ensures that a company is mindful

---

78. *See* Widen, *supra* note 21.

79. IEEE 7000, *supra* note 5, at 9.

80. *Id.*

81. *Id.* at 74. An AV system of interest might be ethically aligned along utilitarian principles which justifies deployment based on a cost-benefit analysis, for example.

82. A safety case is a structured logical argument, supported by evidence, that a system will be acceptably safe. *See, e.g.*, Nancy Leveson, White Paper on the Use of Safety Cases in Certification and Regulation, Aeronautics and Astronautics/Engineering Systems, MIT (undated) (excerpted in J. of Sys. Safety Nov./Dec. 2011), http://sunnyday.mit.edu/SafetyCases.pdf.

of ethical considerations at each stage of product life and identifies all relevant stakeholders.

An important metric of system design for AV technology is how the safety of the new technology compares with the safety of a human driver. As an example of the application of IEEE 7000 to development of AV technology, IEEE 7000's "Transparency Management Process" identifies the ethical value of transparency as requiring the provision of information to all stakeholders (internal and external, short-term and long-term) about how the developer of an AI system has addressed ethical concerns during design.[83] The public, including drivers, pedestrians and cyclists, are relevant stakeholders because the new AV technology impacts their safety and well-being. The ethical standard at a minimum requires disclosure of the deployment standard (if not all the details of how that performance standard will be achieved) and not mere assurances that the technology will be deployed when it is acceptably safe or safe enough.[84]

A. Urgency of the AV Problem

The AV Problem needs an answer now, more so than other ethical issues for AV design raised by the famous "Trolley Problem"[85] or the results of MIT's experimental philosophy poll about "Moral Machines."[86] We face issues similar to the AV Problem now on a smaller scale with current testing of AV technology on our public highways,[87] where high profile fatalities

---

83. IEEE 7000, *supra* note 5, at 49.

84. An interview with Chris Urmson, Aurora's CEO, illustrates the rhetoric of "sufficiently safe" used by industry participants. *See* Jerry Hirsch, *Autonomous Vehicle Pioneer Urmson Talks About Safety and Risks*, TRUCKS (June 22, 2020), https://www.trucks.com/2020/06/22/autonomous-vehicle-risks-urmson/.

85. Judith Jarvis Thomson, *The Trolley Problem*, 94 YALE L.J. 1395 (1985). The name "Trolley Problem" comes from an ethical dilemma where one must make a choice whether or not to pull a lever to direct a trolley onto a track with one worker and away from a track with five, when either choice is fatal to those persons who are hit. It is based on scenarios originally presented by Philippa Foot in 1967. *Id.* at 1395.

86. *See* Edmond Awad et al., *The Moral Machine Experiment*, 563 NATURE 59 (2018).

87. Greg Bensinger, Opinion, *Why Tesla's 'Beta Testing' Puts the Public at Risk*, N.Y. TIMES (July 30, 2021), https://www.nytimes.com/2021/07/30/opinion/self-driving-cars-tesla-elon-musk.html.

involving AVs already have occurred.[88] Moreover, even as many anecdotal accounts of failures of Level 2 technology[89] get headlines, AV companies aim to deploy the more complex Level 4 technology as soon as 2023.[90]

B. Absence of a Clear Deployment Standard Creates Uncertainty

As an example, posit a simple safety rating scale based on number of miles driven without a fatality by an average human driver, expressed on a scale of 1 to 5.[91] Application of this scale illustrates in a simple way the problem caused by the absence of a clear deployment standard. Assume the hypothetical average human driver rates a 3 on this scale. Machine drivers rate a 2 for modest safety improvement over the average human driver, and a 1 for significant improvement in safety. A 4 represents a modest decrease in safety from the average human driver, and 5 a significant decrease in safety.

In principle, Aurora must choose one of two options. Option One: it could keep its deployment standard vague to preserve its deployment options in case financial exigency necessitates a risky premature deployment (less safe than a human driver, at perhaps a 4 or even a 5 on our hypothetical scale). Option Two: it could build trust by announcing that

---

88. *See, e.g.*, Bryan Pietsch, *2 Killed in Driverless Tesla Car Crash, Officials Say*, N.Y. TIMES (Nov. 10, 2021), https://www.nytimes.com/2021/04/18/business/tesla-fatal-crash-texas.html.

89. Tim Levin, *Tesla's Full Self-Driving Tech Keeps Getting Fooled by the Moon, Billboards, and Burger King Signs*, BUS. INSIDER (July 26, 2021, 11:30 PM), https://www.businessinsider.com/sai (use website search feature to find article using the article title).

90. *See, e.g.*, Aurora S-1, *supra* note 1, at 83 (indicating a target deployment date for the trucking industry of late 2023).

91. Any scale used in actual practice needs to address many other details, such as whether the concept of average human driver should exclude impaired persons, and how the road condition of miles driven in testing compares to miles driven as reflected in government statistics. Indeed, a standard of expert human driver would be a better goal than average human driver, and consistent with standards used to test some other automotive systems. *See, e.g.*, INT'L ORG. FOR STANDARDIZATION, ISO 26262-12:2018, annex C.4 (2018) (measuring positive risk balance for motorcycles with reference to expert drivers rather than average drivers).

deployment will only occur after Aurora can justify a safety case that its AV technology rates a 1 for safety.[92]

The standard that an AV technology be "safer than a human driver" (the Safety Proposition) as a condition to initial deployment at scale appears often in the AV discourse. As examples, Daniel Kahneman, a Nobel prize winning behavioral economist, noted that, with respect to AV technology: "[b]eing a lot safer than people is not going to be enough. The factor by which they have to be more safe than humans is really very high."[93] The German Ethics Code states that the primary goal of AV technology ought to be the promotion of safety and an overall positive balance of benefits against burdens.[94] This appears to be the standard that the National Highway Traffic Safety Administration would apply if it produces substantive regulations.[95] New York City's AV regulation uses a "better than a human driver" standard.[96] Though it is not itself a safety standard, J3016:2014 used a "better than a human driver" concept to describe the standard for a high automation system to restore a vehicle to a minimum risk condition as "with at least the level of performance that could be

92. A middling choice of deployment at a 3 rating, or even perhaps a 2 rating, would conflict with public expectations that AV technology will achieve a significant safety improvement and not be merely value neutral. It is not realistic that an AV company would announce a goal of the status quo level of safety.

93. Tim Adams, *Daniel Kahneman: 'Clearly AI Is Going to Win. How People Are Going to Adjust Is a Fascinating Problem'*, THE GUARDIAN (May 16, 2021, 8:00 AM), https://www.theguardian.com/books/2021/may/16/daniel-kahneman-clearly-ai-is-going-to-win-how-people-are-going-to-adjust-is-a-fascinating-problem-thinking-fast-and-slow?msclkid=9174f6e7a6ec11eca0a1d3da029ec840 (reporting observations of Daniel Kahneman).

94. Christoph Luetge, *The German Ethics Code for Automated and Connected Driving*, 30 PHIL. & TECH. 547–58 (2017) (the "German Ethics Code").

95. Framework for Automated Driving System Safety, 85 Fed. Reg. 78,058, 78,060 (proposed Dec. 3, 2020) (to be codified at 45 C.F.R. pt. 571) (noting engineering measures which would seek to show that ADS perform with a "high level of proficiency").

96. "New York City is implementing a permit process, including self-certifications from autonomous vehicle technology companies that their autonomous vehicles will operate more safely than human drivers in New York City . . . ." N.Y.C., N.Y., RULES OF THE CITY OF NEW YORK tit. 34, ch. 4, § 4-17 (2021).

expected from a *human driver* under the same conditions."[97] Variations of the Safety Proposition appear in many corners of the AV discourse.

The Aurora S-1 even notes this better-than-a-human-driver performance standard as a risk factor: "[t]he industry can be characterized by a significant number of technical and commercial challenges, including an expectation for better-than-a-human driving performance . . . ."[98] But rather than stating its own deployment standard as better-than-a-human driving performance, Aurora states its goal as "achieving sufficiently safe self-driving system performance as determined by us, government & regulatory agencies, our partners, customers, and the general public."[99]

The timing of application of the deployment standard, which Aurora leaves opaque, is critical. Per IEEE 7000, "[o]rganizations that do not explicitly define their ethical values are more likely to encounter ethical issues, such as placing economic gain or privileges of a few above human rights . . . ."[100] If Aurora publicly adopted a safety rating of 1, it would be less likely to deploy at a rating of 4 or 5 when facing a financial exigency. A deployment at a rating of 4 or 5 justified by the expectation of future benefits might prove controversial, if not ethically questionable.

Option One preserves a harm now, benefits later utilitarian justification[101] for early deployment of vehicles with high automation

---

97. SAE INT'L, TAXONOMY AND DEFINITIONS FOR TERMS RELATED TO ON-ROAD MOTOR VEHICLE AUTOMATED DRIVING SYSTEMS J3016_201401 10 (2014), https://www.sae.org/standards/content/j3016_201401/.

98. Aurora S-1, *supra* note 1, at 7.

99. *Id*. A standard of "sufficiently safe" allows for lobbying efforts to convince regulators to allow deployment on a harm now, benefits later analysis—a justification which, if disclosed, might cause public outrage.

100. IEEE 7000, *supra* note 5, at 73 annex H. We do not suggest that deployment at a level less than a rating of 1 would violate a human right. IEEE 7000 uses violation of a human right as illustrative of a negative consequence of failure to explicitly define ethical values. *Id*.

101. IEEE 7000 specifically identifies utilitarianism as an "[e]thical decision-making approach to consider the consequences of system design *and deployment* (harms and benefits)." *Id.* at 22 (emphasis added). Section 5.6 includes utilitarian ethics as one of three used in the standard to help identify and prioritize values in accordance with the standard. *Id.* at 30. The standard notes that general utilitarian ethics considers "the

technology at an SAE rating of Level 3, 4, or 5 that is less safe than a human driver in the near term but with the expectation that the technology will become safer than a human driver in the long run. This is a classic trade-off identified by IEEE 7000 as "[d]ecision-making actions that select from various requirements and alternative solutions on the basis of net benefits to the stakeholders."[102] If Aurora wants to elect Option One, IEEE 7000 requires public disclosure now in accord with its recommended value of transparency so that an informed public debate might begin. The risk for Aurora is that the public might not readily accept such a harm now, benefits later justification for deployment. Moreover, there is no assurance that the future benefits will materialize—further complicating any utilitarian analysis.

IEEE 7000, however, values transparency, which includes transfer of information to a stakeholder (here, the public)[103] and indicates the social responsibility of an organization is an "[o]bligation to wider society to respect the values reigning within it."[104] If the public as a stakeholder has an interest that any deployment at scale of AV technology only occur if it is safer than a human driver at the time of deployment, respect for this value requires transparency in the form of disclosure, particularly if an organization intends to go in a different direction. Indeed, without this disclosure, it is difficult to even determine the public's appetite for a harm now, benefits later approach.

Consistent with IEEE 7000, Option Two builds public trust if Aurora's management must defend a safety case to rate its AV technology a 1 to an independent committee of its board of directors before deployment.[105]

---

consequences for both direct and indirect stakeholders in the short, middle, and long terms." *Id.* at 58 annex C.

102. *Id.* at 22. IEEE 7000 is very clear that "society at large" and the "general public" are considered stakeholders. *Id*. § 5.4, at 27.

103. *Id.* at 22.

104. *Id.* at 21.

105. Aurora has formed a Safety Advisory Board. However, there has been no public commitment to grant that board veto power on a deployment decision the Safety Board considers insufficiently safe. *See* Nat Beuse, *Our Updated Safety Report and First-Ever Safety Advisory Board*, AURORA (June 2, 2021), https://aurora.tech/blog/aurora-shares-safety-report (failing to describe the powers of the Safety Advisory Board). The safety advisory board is not referenced in any of: Aurora Innovation, Inc., Certificate of Incorporation

(IEEE 7000 recommends appointing project team members to various roles to support value-based engineering efforts,[106] though it does not require engagement of an ethics expert to conform to the standard.)[107] Adopting corporate governance structures to protect the integrity of deployment decisions, combined with a robust corporate ethics code, would work with IEEE 7000 to strengthen a commitment to deploy only when evidence justifies a claim that an overall safety improvement immediately follows deployment.[108]

Disclosing deployment standards in SEC filings provides additional practical incentives against making safety a secondary concern in the face of financial exigency.[109] Following IEEE 7000 makes an even stronger case for trust while simultaneously conforming to a new industry standard.

(Form 8-K, Exhibit 3.1) (Nov. 3, 2021) https://www.sec.gov/Archives/edgar/data/0001828108/000119312521319524/d249044dex31.htm; Aurora Innovation, Inc., Bylaws (Form 8-K, Exhibit 3.2) (Nov. 3, 2021) https://www.sec.gov/Archives/edgar/data/0001828108/000119312521319524/d249044dex32.htm; *Code of Conduct and Ethics*, AURORA, https://d1io3yog0oux5.cloudfront.net/_55e9196521c7b242e94ca84f97b8d5e7/aurora/db/867/7487/file/Code+of+Conduct+and+Ethics.pdf (last visited April 12, 2022); or *Corporate Governance Guidelines*, AURORA, https://d1io3yog0oux5.cloudfront.net/_54cafb1a3202757f72937b6681f4bf93/aurora/db/867/7488/file/Aurora+Corporate+Governance+Guidelines.pdf (last visited April 12, 2022). Given the lack of any formal role for the Safety Advisory Board, it is difficult to see how it might operate as a substitute for compliance with IEEE 7000.

106. IEEE 7000, *supra* note 5, at 32–35. These roles include a "Value Lead" who bridges the gap between engineering, management, and ethical values in a constructive way, a "User Advocate" who represents the direct and indirect users of the system, and a "Transparency Manager" who leads the communication of technical decisions and system functions to stakeholders. A "System Expert" has the responsibility to listen to stakeholders. *Id.* at 33. Aurora's Safety Advisor Board does not perform this function as it does not appear to assign actual project team members within the company to perform these important roles. *See* Beuse, *supra* note 105.

107. *Id.* at 26.

108. The standard does not purport to specify ethical requirements for non-engineering areas of organizational governance and ethical policies. It only applies to structures which directly affect the design of a system of interest. *Id.*

109. Indeed, some in the AV industry have suggested a "Safety Third" attitude which, even if made in jest, raises concerns. *See* Max Chafkin & Mark Bergen, *Fury Road: Did Uber Steal the Driverless Future From Google?*, BLOOMBERG (Mar. 16, 2017, 1:00 AM),

The AV industry's mantra to date, as exemplified by the Aurora S-1, can be summarized: just trust us, we are smart, we will do the right thing.[110] When pressed, the AV industry references vague content-free standards such as "sufficiently safe."[111] But announcing a meaningful deployment standard and supporting that standard with deployment decision procedures builds trust more effectively than naked appeals to trust—without publicized standards for deployment and protective corporate governance structures, a stronger case for regulation exists.

The AV industry resists regulation, arguing that regulations will become outdated before becoming operational, slowing technological progress,[112] while simultaneously arguing a utilitarian case for early

https://www.bloomberg.com/news/features/2017-03-16/fury-road-did-uber-steal-the-driverless-future-from-google (noting that "[s]omeone [at Otto, an AV company,] had distributed stickers—in OSHA orange—with a tongue-in-cheek slogan: 'Safety third.'").

110. The Aurora S-1 advances this message by assuring investors that Aurora does not hire "jerks." Aurora S-1, *supra* note 1, at 84.

111. *See* Aurora S-1, *supra* note 1, at 7; *see also* Hirsch, *supra* note 84. A Rand study explores the concept of "sufficiently safe" and how that determination might be made. *See* MARJORY S. BLUMENTHAL ET AL., RAND CORP., SAFE ENOUGH: APPROACHES TO ASSESSING ACCEPTABLE SAFETY FOR AUTOMATED VEHICLES (2020), https://www.rand.org/content/dam/rand/pubs/research_reports/RRA500/RRA569-1/RAND_RRA569-1.pdf.

112. *See* ISSUES IN AV TESTING, *supra* note 2, at 8.

deployment.[113] Yet, the AV industry's recent BEST PRACTICE STATEMENT[114] reveals no statistically significant metrics, standards, or data to back up any utilitarian claim that current AV technology reduces highway fatalities or that it will do so in the future. This violates the IEEE 7000 requirement of verification, which demands "[c]onfirmation, through the provision of objective evidence, that specified requirements have been fulfilled."[115]

The AV industry currently does not have an objective method to assess the safety of AV technology relative to that of a human driver. Instead, the AV industry conducts a mere public relations campaign using PAVE, its 501(c)(3) tax exempt organization, to convince the public of AV technology's potential benefits.[116] However, IEEE 7000 does not recognize public relations efforts as relevant to ethical AI design (apart from a commitment to transparency).

---

113. Chris Urmson, Aurora's CEO, echoes utilitarian justifications for rapid deployment of AV technology when he suggests that delays in implementation of AV technology will cost lives.

> One of the parts that maybe gets a little bit lost is that we need to be careful and thoughtful about what the threshold is that we accept of risk. We obviously want to drive that to zero over time. But it's very easy to overlook the fact that the status quo is broken. There's an incredible opportunity to move from the status quo towards zero. We should be saving those lives along the way and not wait for the perfect at the expense of all those lives.

Hirsch, *supra* note 84. This quotation is highly suggestive of urging a harm now, benefits later justification for deployment of AV technology that, at the time of initial deployment, is less safe than the average human driver.

114. SAE INT'L, AVSC BEST PRACTICE FOR METRICS AND METHODS FOR ASSESSING SAFETY PERFORMANCE OF AUTOMATED DRIVING SYSTEMS (ADS) (2021) [hereinafter BEST PRACTICE STATEMENT], https://www.sae.org/standards/content/avsc00006202103/.

115. IEEE 7000, *supra* note 5, at 23.

116. Individual automakers supplement this coordinated advertising campaign with their own efforts. *See, e.g.*, *Path to Autonomous*, GENERAL MOTORS, https://www.gm.com/commitments/path-to-autonomous.html (last visited Mar. 18, 2022) (profiling a chief AV engineer as a "mother of three children who will be driving soon, . . . motivated by the role AVs play in GM's vision of a world with Zero Crashes, Zero Emissions and Zero Congestion"). GM hopes that referencing the support of a "mother of three" will lead the public to conclude that AV technology is safe. One can only conclude, however, that the mother of three is an engineer employed by GM who believes in AV technology's potential.

It is ironic that the Aurora S-1 places the goal of building "trust" as a centerpiece of its business strategy yet makes purely hypothetical utilitarian calculations concerning relative public safety without meaningful data. Indeed, Aurora states its philosophy as "build and earn trust with everything that we do."[117] Aurora's rhetoric aligns with the messaging of PAVE. The AV industry recognizes that the public is wary of self-driving technology safety, particularly as the industry increases testing on public highways. Yet, Aurora takes care to make no commitment to a standard for deployment of AV technology at scale to assure a concerned public. Aurora apparently wants the flexibility to make a harm now, benefits later justification without identifying that it is preserving this option.

C. The Moral Hazard

The presence of a moral hazard in the initial deployment decision is corrosive of trust, further highlighting the importance of following IEEE 7000. A moral hazard exists because the corporate form used to operate Aurora's business shields investors and management from personal liability for the consequences of any mistaken decision by Aurora to deploy AV technology at scale before it is safe to do so.[118] When Aurora must decide whether to deploy or delay for more development and testing, its management will face enormous financial pressure to deploy. The Aurora S-1 suggests Aurora will have a market value of over $10 billion, though it currently loses money, and will continue to lose money in the near term.[119] The Reinvent S-4 prepared for the shareholder vote prior to Aurora's IPO indicated no positive EBITDA until projected free cash flow materializes in 2027.[120] A reasonable assumption on the available financial information presented is that, if deployment is delayed, Aurora will fail.[121]

---

117. *See* Aurora S-1, *supra* note 1, at 84. A better approach to building trust would answer the deployment question directly, rather than populating the Aurora website with volumes of essentially content free praise for a safety culture.

118. By "mistaken" decision, we mean "mistaken" from the vantage point of maximizing social welfare. The moral hazards caused by the limited liability associated with corporations is well known. *See, e.g.*, William H. Widen, *Corporate Form and Substantive Consolidation*, 75 GEO. WASH. L. REV. 237 (2007).

119. Aurora S-1, *supra* note 1, at 7.

120. Reinvent S-4, *supra* note 1, at 30, 130–32.

121. For securities law liability reasons, projections may appear in a registration statement on Form S-4 but not in a registration statement on Form S-1. There are two

Aurora might take four different stances towards the potential moral hazard. If Aurora picks Choice One, it can either tell the public that deployment might occur when the Safety Proposition is false, or it might remain silent. Disclosing its true stance towards safety in the case of Choice One creates a serious public relations risk. If Aurora does not amend the Aurora S-1 and remains silent, that is a good indication that Aurora is preserving the option to deploy when it either has no idea about the truth of the Safety Proposition or it has reason to believe it is false.[122] This is so because if Aurora's ethical values and principles allow it to deploy its AV technology at scale when the Safety Proposition is false or its status unknown, then its acceptance of this possibility will not conflict with the financial interests of its investors.

This is a hazard for the public, but not a moral hazard for Aurora's management in the classic sense; this attitude towards safety will never conflict with a fiduciary duty to stockholders because preserving an option for stockholders always has value. It is always better to choose an option which *may* not result in the loss of $10 billion by deploying early, than *accept* the certain loss of $10 billion resulting from failure to deliver a product on time.

If Aurora picks Choice Two and adopts the moral principle that it will not deploy AV technology when the Safety Proposition is false (or when its truth or falsity is unknown), then it might make express disclosure of its

---

other alternatives to financial failure. If Aurora can develop its AV technology sufficiently to demonstrate a "proof of concept," then another industry buyer might acquire Aurora to obtain its technology, or the proof of concept might be enough to secure another round of financing. But, as a stand-alone company, Aurora likely fails if deployment is delayed in any material way. An effort to show proof of concept focuses on demonstrating functionality and not safety, assuming any accidents during testing can be paid for and subsequently advertised as "fixed." The cost of a few lives may not provide an adequate deterrent given the monetary stakes. By comparison, Embark was demonstrating a proof of concept when it sent a truck on an autonomous journey around Oakland. *See* Embark S-4, *supra* note 33.

122. One way for Aurora to announce a decision to opt for Choice Two post public offering, without amending the Aurora S-1 or Reinvent S-4, would be to make a corporate decision to implement IEEE 7000 and, as part of that implementation decision, announce a standard for initial deployment. As a new development this decision would be reported on an SEC Form 8-K—avoiding the appearance that a prior filing contained a material misstatement or omission that needed correction.

stringent principle for deployment or it might remain silent. The option to remain silent having made Choice Two makes no sense if the goal is to build trust because it is a missed opportunity to create the public trust which Aurora strives to achieve.

The failure to clearly identify a standard for deployment of AV technology at scale, coupled with a failure to comply with IEEE 7000, is corrosive of trust. Indeed, lack of clarity about the deployment standard and failure to implement IEEE 7000 may conflict with undertakings made by IEEE members. Among other things, IEEE members agree "to strive to comply with ethical design," "to avoid real or perceived conflicts of interest," and "to be honest, and realistic in stating claims or estimates based on available data."[123] IEEE members who serve as officers, directors, or advisors to AV companies need to justify a failure to implement IEEE 7000, which is designed to fulfill the agreements each member commits to every year.

III. THE DIRTY DOZEN MYTHS ABOUT AV TECHNOLOGY

AV testing on public roads poses serious risks to vulnerable road users. Despite these risks, the AV industry campaigns for favorable regulatory treatment for both current testing and future general deployment. This campaign to limit meaningful regulation employs various myths about AVs which are easily debunked (as demonstrated below). The industry's use of these myths is inconsistent with the development of trustworthy AI.

The *Trustworthy AI Guidelines* emphasize the "freedom of the individual": "[i]n an AI context, freedom of the individual for instance requires mitigation of (in)direct illegitimate coercion . . . deception and unfair manipulation."[124] The *Guidelines* also strive to preserve "human dignity": "[i]n this [AI] context, respect for human dignity entails that all people are treated with respect due to them as moral *subjects*, rather than merely as *objects* to be sifted, sorted, scored, herded, conditioned or

123. IEEE, IEEE CODE OF ETHICS (2020), https://www.ieee.org/content/dam/ieee-org/ieee/web/org/about/corporate/ieee-code-of-ethics.pdf.

124. *Trustworthy AI Guidelines*, *supra* note 6, at 10.

manipulated."[125] The *Guidelines* further emphasize that AI systems should be developed in a manner that respects human dignity.[126]

Emphasis on the ethical use of AI is commonly focused on direct concerns, such as dodges that attempt to use human drivers as a moral crumple zone when deploying unreliable AI features.[127] However, indirect concerns are just as important, such as the legal and regulatory environment in which the AI technology is developed.

For the purposes of this analysis, an indirect form of coercion, deception, or manipulation occurs when advocacy and talking points used to shape the legal environment contain untruths, half-truths, and substantive omissions. The AV industry has drifted into the realm of propaganda by use of the dirty dozen myths described (and debunked) below to shape the legal regime in which AV technology is developed. This provides an additional reason for the public to withhold trust from the AV industry. At the level of ordinary ethical intuition, consider the following list of myths and our analysis to decide whether a person using these myths is worthy of trust.[128]

---

125. *Id.*

126. *Id.*

127. A "moral crumple zone" approach to analysis of an accident is one that employs a human to absorb moral and legal consequences when machinery malfunctions. *See* Madeleine Clare Elish, *Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction*, 5 ENGAGING SCI., TECH., & SOC'Y 40 (2019) (describing how responsibility for an action may be misattributed to a human actor who had limited control over the behavior of an automated or autonomous system).

128. One of the authors published an earlier version of this list of myths. Philip Koopman, *Autonomous Vehicle Myths: The Dirty Dozen*, EE TIMES (Oct. 22, 2021), https://www.eetimes.com/autonomous-vehicle-myths-the-dirty-dozen/.

MYTH #1: 94 PERCENT OF CRASHES ARE DUE TO HUMAN DRIVER ERROR, SO AVS WILL BE SAFER.

The informal version of this myth is that humans drive drunk or fall asleep or text while driving. Therefore, computer drivers will necessarily be safer than human drivers.

To be sure, many crashes are caused by impaired drivers. However, the 94 percent figure is a misrepresentation of the original source.[129] In fact, the National Highway Traffic Safety Administration (NHTSA) study data shows only that in 94 percent of crashes, a human driver might have helped avoid a bad outcome. That is not the same as causing a crash. Indeed, the source explicitly disavows placing 94 percent "blame" on the human driver:

> The critical reason [generally, the last event in the causal chain of the crash] was assigned to drivers in an estimated 2,046,000 crashes that comprise 94 percent of the NMVCCS crashes at the national level. However, in none of these cases was the assignment intended to blame the driver for causing the crash.[130]

Not only does 94 percent not represent simple driver error, it elides the fact that AVs will make different kinds of mistakes than human drivers.[131] This myth is particularly troublesome precisely because the 94 percent does not describe driver error. Wrongly treating that number as representing driver error downplays the need to watch for AV errors. To be sure, AV technology will improve over time, but it remains to be seen how long it will take for AVs to be net safer than human drivers in complex driving situations after factoring AV's shortcomings into the analysis.

---

129. *See, e.g.*, Don Kostelec, *The 94% Error: We Need to Understand the True Cause of Crashes*, STREETSBLOG USA (Oct. 14, 2020), https://usa.streetsblog.org/2020/10/14/the-94-solution-we-need-to-understand-the-causes-of-crashes/.

130. SANTOKH SINGH, NAT'L HIGHWAY TRAFFIC SAFETY ADMIN., REP. NO. DOT HS 812 115, CRITICAL REASONS FOR CRASHES INVESTIGATED IN THE NATIONAL MOTOR VEHICLE CRASH CAUSATION SURVEY 1 (2015), https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812115.

131. One example is a Tesla vehicle mistaking a moon near the horizon for a yellow traffic signal lamp. *See* Jay Ramey, *Tesla's Full Self Driving System Mistakes the Moon(!) for Yellow Traffic Light*, MSN (July 23, 2021), https://www.msn.com/en-us/autos/news/tesla-s-full-self-driving-system-mistakes-the-moon-for-yellow-traffic-light/ar-AAMuszU.

MYTH #2: YOU CAN HAVE EITHER INNOVATION OR REGULATION—NOT BOTH.

Car makers and their representatives encourage removal of "regulatory barriers to AV deployment," and warn that prescriptive requirements of "a specific approach . . . could stifle innovation."[132] Industry talking points in various venues seek to create an expectation that regulation necessarily impedes innovation.

This is a false dilemma because regulation need not impede innovation.[133] One way to avoid impeding innovation is to adopt regulations which merely require the industry to follow its own design and operational safety standards rather than setting specific technology-based regulatory test regimes. For example, regulators could avoid setting detailed technical requirements for road testing safety themselves, and instead require conformance to the SAE J3018 standard.[134] That standard helps ensure that the human safety driver is properly qualified and trained. It also requires that testing be done in a responsible manner, consistent with good engineering validation and road safety practices. It places no constraints on the AV technology being tested other than requiring a means for a trained human test driver to take over immediate control of the test vehicle to intervene when required to maintain safety.[135]

None of the regulatory standards proposed by NHTSA[136] stifle innovation. Rather, they promote a level playing field so that companies

---

132. All. for Auto. Innovation, Comment Letter dated April 1, 2021 on Proposed Rule of the NHTSA Framework for Automated Driving Systems (Dec. 3, 2020), https://www.regulations.gov/comment/NHTSA-2020-0106-0728.

133. For a discussion of the status of state and local regulation of AV technology, see Joshua Burd, *State Regulation Fosters Autonomous Vehicle Development*, THE REGUL. REV. (Oct. 14, 2021), https://www.theregreview.org/2021/10/14/burd-state-regulation-fosters-autonomous-vehicle-development/ (noting that state and local regulation often is very permissive but sometimes this flexibility compromises safety).

134. J3018, *supra* note 28.

135. *Id.* SAE J3018 presumes the safety driver will be physically present in the test vehicle as a matter of scope. *Id.* It does not prohibit testing with a remote safety driver. *Id.* In a standards-based regulatory regime, regulators would ask the industry to expand the scope of that standard to include any adjustments appropriate to remote safety drivers.

136. The proposed standards are primarily ISO 26262, ISO 21448 and ANSI/UL 4600. *See* NHTSA Framework for Automated Driving System Safety, 85 Fed. Reg. 78058, 78065-66

cannot skimp on safety to gain competitive advantage while putting other road users at undue risk.

If a company adopts safety as its first priority, as many say they do, there is no reason to believe that they cannot also comply with a regulatory mandate to follow industry-consensus safety standards. Such standards are written and approved by the industry itself via an accredited Standards Development Organization (SDO) such as SAE International, the International Standards Organization (ISO), the American National Standards Institute (ANSI), or Underwriters Laboratories. The SDO process requires rigorous review from stakeholders, including voting representatives of AV developers.

As a normative matter, the AV industry ought to compete on features other than safe operation of an AV system—with acceptably safe operation of an AV system held to a uniformly very high standard, driven by SDO-created industry standards (as has been the practice in aviation, rail, chemical processing, and other life critical applications of computer technology for decades).[137]

MYTH #3: THERE ARE ALREADY SUFFICIENT REGULATIONS IN PLACE.

A claim that sufficient regulations are already in place is sometimes made directly by AV industry participants conducting tests but often takes the more subtle form of saying that a particular AV industry tester conforms to all existing regulations.[138]

(proposed Dec. 3, 2020) (to be codified at 49 C.F.R. pt. 571) [hereinafter NHTSA DEC. 3, 2020 ANPRM], https://www.govinfo.gov/content/pkg/FR-2020-12-03/pdf/2020-25930.pdf.

137. Competing on the safe operation of an AV system differs in kind from competing on other safety features, such as performance in crash tests, because these safety features are add-ons and enhancements, not replacement of a human driver. Historical competition on safety and so-called "star rating systems" primarily relate to mitigation of crash consequences on the assumption that there will be an imperfect human driver. The primary safety argument in favor of an AV is that its computer system will be a better driver than a human insofar as the AV will not crash in the first place. Those aspects of a system which replace a human driver, and especially the software aspects of such a system, ought to be uniformly high and not subject to traditional competition for other aspects of vehicles.

138. The authors have heard this argument made by AV company advocates in verbal exchanges that are not formally citable.

Some states, such as Texas and Arizona, enforce no practical limitations on AV testing so it is particularly easy to conform to all existing regulations. Other states, such as Pennsylvania and California, require registration and some form of reporting.[139] But no state requires adherence to a safety standard relevant to AVs. The one exception currently is New York City, which requires conformance to SAE J3018 for public road testing.[140]

Thus, regulatory assurance of safety, if any, is little more than taking the manufacturer's word for it. More is required.

MYTH #4: WE DON'T NEED PROACTIVE AV REGULATION BECAUSE OF EXISTING REGULATIONS AND PRESSURE FROM LIABILITY EXPOSURE.

The current Federal Motor Vehicle Safety Standards (FMVSS)[141] do not cover computer-based system safety. They are primarily about testing headlights, seat belts, air bags, and other basic non-AV vehicle safety functions. An AV that complies with the FMVSS, while having passed useful and important tests, is not necessarily acceptably safe (for example, free of unreasonable risk) for use on roadways even as a conventional vehicle, let alone as an AV.

The National Highway Traffic Safety Administration (NHTSA) generally operates reactively to bad events. If car companies do not voluntarily disclose issues, many injury and fatality loss events are typically required before NHTSA forces action. For example, it took eleven crashes involving Teslas on "autopilot" colliding with emergency vehicles over 3.5 years to prompt NHTSA action.[142]

Safety regulators should think hard about an approach in which "safety" means requiring insurance to compensate the next of kin after a fatality, which is the typical requirement imposed by state regulations. With multi-billion-dollar development war chests, a few million dollars of payout after a mishap seems scant deterrent to safety shortcuts in the race to autonomy.

---

139. *See, e.g.*, *Autonomous Vehicles | Self-Driving Vehicles Enacted Legislation*, NAT'L CONF. OF STATE LEG. (Feb. 18, 2020), https://www.ncsl.org/research/transportation/autonomous-vehicles-self-driving-vehicles-enacted-legislation.aspx.

140. N.Y.C., N.Y. RULES OF THE CITY OF NEW YORK tit. 34, ch. 4, § 4-17 (2021).

141. NHTSA Federal Motor Vehicle Safety Standards, 49 C.F.R. § 571, Subpart B (1971) (omitting references to computer-based system safety).

142. NHTSA, ODI RESUME: INVESTIGATION PE 21-020 (2021), https://static.nhtsa.gov/odi/inv/2021/INOA-PE21020-1893.PDF.

MYTH #5: EXISTING SAFETY STANDARDS ARE NOT APPROPRIATE BECAUSE (PICK ONE OR MORE):

- they are not a *perfect* fit;
- no *single* standard applies to the whole vehicle;
- they would reduce safety because they prevent the developer from doing more;
- they would force the AV to be less safe;
- they were not written *specifically* for AVs.

Voluntary Safety Self-Assessment reports (VSSAs) issued by AV companies commonly assert variations on these themes to argue that industry standards somehow do not apply to very special, unique AV technology. For example, the Waymo safety methodology report issued as a supplement to their VSSA—which lay readers might tend to interpret as a pro-standard approach—does not actually commit Waymo to following any relevant AV safety standard.[143] Other VSSA documents simply roll call standards while making no commitment to adhere to them.[144]

These statements misrepresent how actual safety standards work. ISO 26262, ISO 21448, and ANSI/UL 4600 all permit significant flexibility in support of safety. All three work together to fit any safe AV.

ISO 26262[145] ensures safe operation for conventional computer-based functions. ISO 21448[146] deals with the inherent limitations in sensors, and surprises in an open external environment, by covering so-called Safety of

143. *See* N. Webb, et al., *Waymo's Safety Methodologies and Safety Readiness Determinations* 6–9, WAYMO LLC (Oct. 31, 2020), available at *Safety Publications*, *Safety*, WAYMO, https://waymo.com/safety/ (last visited Mar. 18, 2022) (follow "Safety Methodologies" hyperlink).

144. *See, e.g.*, AURORA, VOLUNTARY SAFETY SELF-ASSESSMENT 12 (2021), available at *Aurora Voluntary Safety Self-Assessment*, AURORA, https://aurora.tech/vssa (last visited Mar. 18, 2022) (follow "View Aurora's VSSA" hyperlink).

145. Int'l Org. for Standardization [ISO], *ISO 26262-1:2018 Road vehicles — Functional safety* (2018) [hereinafter ISO 26262].

146. ISO, *ISO/DIS 21448 Road vehicles — Safety of the intended functionality* (undergoing final revision process prior to issuance).

The Intended Functionality (SOTIF) for AVs. ANSI/UL 4600[147] works with ISO 21448 and ISO 26262 to cover AV system-level safety, encompassing the vehicle and its support infrastructure.

The US Department of Transportation (US DOT) has already proposed this set of standards in an Advanced Notice of Proposed Rulemaking.[148] All these standards allow developers to do more than required and are flexible enough to accommodate any AV. None force a company to be less safe (a truly strained argument to criticize standards largely drafted by industry participants who might complain about following standards). None constrain the technical autonomy approach beyond requiring safety.

MYTH #6: LOCAL AND STATE REGULATIONS NEED TO BE STOPPED TO AVOID A "PATCHWORK" OF REGULATIONS THAT INHIBITS INNOVATION.

The industry and Federal Government frequently bemoan the threat of a "patchwork of incompatible laws"[149] for AV safety.

A significant reason that local and state regulations are developing as a "patchwork" approach is that in each jurisdiction, the AV companies play hardball, negotiating to minimize regulation. The companies threaten that essentially any fettering of testing with safety regulations will create "one of the least hospitable cities in the US for AV development," for example, calling upon regulators to "reject these additional hurdles to New York's autonomous vehicle future."[150] The typical playbook for the AV industry (as reflected in off the record remarks by some elected officials) is to threaten to take jobs and spending elsewhere if there is substantive safety

147. UNDERWRITERS LAB'YS, ANSI/UL 4600 STANDARD FOR SAFETY FOR THE EVALUATION OF AUTONOMOUS PRODUCTS (1st ed. 2020).

148. *See* NHTSA DEC. 3, 2020 ANPRM, *supra* note 136.

149. *See, e.g.*, ISSUES IN AV TESTING, *supra* note 2; NHTSA, FEDERAL AUTOMATED VEHICLES POLICY 7 (2016) (noting the objective of a consistent national framework rather than a patchwork of incompatible laws), https://www.nhtsa.gov/sites/nhtsa.gov/files/federal_automated_vehicles_policy.pdf.

150. *See, e.g.*, Gersh Kuntzman, *Self-Driving Car Industry, Promising Safety, Pushes Back on DOT Plan to Regulate Testing*, STREETSBLOG NYC (Sept. 1, 2021), https://nyc.streetsblog.org/2021/09/01/self-driving-car-industry-promising-safety-pushes-back-on-dot-plan-to-regulate-testing/ (describing an NYC DOT open meeting).

regulation, as well as threaten the area with a reputation for being hostile to innovation and technology.

The outcome of each negotiation is different, resulting in somewhat different regulations or voluntary guidance. In truth, the patchwork is largely self-inflicted by the AV companies themselves.

Moving to regulation based on industry standards would help the situation by establishing a level playing field across all states and municipalities. A federal regulation that prevents states from acting but does not itself ensure safety would make things worse.

MYTH #7: WE CONFORM TO THE "SPIRIT" OF SOME STANDARD.

Typically, the "spirit" statements made by AV developers rely on the notion that there might be a need for deviation from the standard.[151] Yet a concrete example of such a need for deviation is never really stated, nor do the AV developers elaborate in any concrete way what it might mean to conform to the "spirit"—as opposed to conforming to both the spirit and the letter of the standard.

The industry promulgated standards are all flexible enough that if a company conforms to the spirit of the standard in a meaningful way, it can readily conform to the letter of the standard as well. However, if a company is in a hurry or wants to cut costs, committing to follow only the spirit might come in handy as a form of evading any expectation of following industry safety standards. A better practice would involve consultation with regulators to either confirm the reasonableness of any required deviation from a standard or obtain a limited exemption from industry safety

151. The term "spirit" is commonly used by industry participants in verbal discussions that are not readily citable. However, equivalent sentiments are seen in Waymo's use of the phrase "informed by … existing safety standards," WAYMO, WAYMO SAFETY REPORT 11 (2021), https://storage.googleapis.com/waymo-uploads/files/documents/safety/2021-12-waymo-safety-report.pdf; Uber's statement that it "follow[s] an internal process informed by the principles of relevant industry standards," UBER ADVANCED TECHS. GRP., A PRINCIPLED APPROACH TO SAFETY 47 (2020), https://uber.app.box.com/v/UberATGSafetyReport; and Ford's statement that its "functional safety process is strongly aligned with the industry automotive safety standard (ISO 26262)," FORD, A MATTER OF TRUST 2.0, FORD'S APPROACH TO DEVELOPING SELF-DRIVING VEHICLES 26 (updated June 2021), https://media.ford.com/content/dam/fordmedia/North%20America/US/2021/06/17/ford-safety-report.pdf.

standard conformance that is properly structured to preserve safety—if, indeed, any such need can be established.

Companies that truly value safety should support transparent conformance to industry consensus standards to raise the bar for competitors. If they don't, that could provide protective cover for any potential bad actors to make hollow claims related to standards that amount to safety theater.

Consider whether a passenger would wish to ride in an autonomous airplane whose manufacturer said: "We conform to the spirit of the aviation safety standards, but we're very smart and our airplane is very special, so we took liberties. We make no concrete claim at all as to standards conformance and involved no independent safety oversight. Trust us; everything will be fine."

MYTH #8: GOVERNMENT REGULATORS AREN'T SMART ENOUGH ABOUT THE TECHNOLOGY TO REGULATE IT.

The proposed US DOT plan to invoke industry standards mentioned above in Myth #5 makes sense because it addresses this concern directly. Industry has already created relevant safety standards. Regulators can simply say: "follow your own industry safety standards, just like all the other safety critical industries do."

If we could trust any industry to self-police safety in the face of short-term profit incentive and inevitable organizational dysfunction, we wouldn't need regulators. But that isn't the real world. Achieving a healthy balance between the industry taking responsibility for safety and oversight from regulators is important.

MYTH #9: DISCLOSING TESTING DATA GIVES AWAY THE SECRET SAUCE FOR AUTONOMY.

Road testing safety is all about whether a human safety driver can effectively keep a test vehicle from creating elevated risk for other road users. That has nothing to do with autonomy-related intellectual property, the point of which is to dispense with the human safety driver after testing has been completed.

Testing safety data need not include anything about the autonomy design or functional performance. For example, consider reporting how often test drivers fall asleep while testing. A non-zero result might be embarrassing (and indicate some level of risk to road users that should be mitigated further), but how does that divulge secret autonomy technology data?

Metrics derived from consistency of conformance to processes in SAE J3018 should provide a way to measure the effectiveness of road-testing safety processes. Such an approach would create measurements for drivers and test protocols, not the underlying technology.

MYTH #10: DELAYING DEPLOYMENT OF AVS IS TANTAMOUNT TO KILLING PEOPLE.

The safety benefits of AVs are aspirational, promised at some ever-receding horizon in the future.[152] Moreover, there is no real proof to show that AVs will be substantially safer than human-driven vehicles, especially when competing with active safety features for human-driven vehicles such as automated emergency braking.[153]

Ignoring industry best practices and putting vulnerable road users at risk today in a bid to maybe, perhaps, someday, eventually avoid future harm if the technology proves economically viable should not be permitted.

Further, bad press from a high-profile mishap can easily set the whole industry back. No company should be rolling the safety-shortcut dice to hit a near-term funding milestone while risking both people's lives and the

---

152. Elon Musk of Tesla is famous for suggesting that functional and safe autonomy is imminent, though the dates he predicts for achieving this goal repeatedly get pushed into the future. *See* Aarian Marshall, *Elon Musk Promises a Really Truly Self-Driving Tesla in 2020*, WIRED (Feb. 19, 2019, 8:53 PM), https://www.wired.com/story/elon-musk-tesla-full-self-driving-2019-2020-promise/# (noting different times Musk missed target dates for deployment of AV technology). *Cf*. Hirsch, *supra* note 84 (quoting Aurora's Urmson who urges early deployment of AV technology now, or in the near future, before it has achieved its promise of zero accidents).

153. *See* JAMES M. ANDERSON ET AL., RAND CORP., AUTONOMOUS VEHICLE TECHNOLOGY: A GUIDE FOR POLICYMAKERS 154 (2016), https://www.rand.org/content/dam/rand/pubs/research_reports/RR400/RR443-2/RAND_RR443-2.pdf (stating that a Tier 1 executive of original equipment manufacturers saw AV technologies as "comparable to 'today's active safety warning systems'" in safety). The most that can be said is that it *seems likely* safety will improve: "AV technology will likely lead to substantial reductions in crashes and the resulting human toll." *Id.* at 16. However, even this analysis is subjective and not based on any data supporting the proposition that safety improvement will be realized within any defined timeline.

reputation of the entire industry. And yet, it seems that AV companies are heavily incentivized to do this very thing.

MYTH #11: WE HAVEN'T KILLED ANYONE, SO THAT MUST MEAN WE ARE SAFE.

In other words: "we've gotten lucky so far, so we plan to get lucky in the future." If there is no evidence of robust, systematic safety engineering and operational safety practices, this amounts to a gambler on a winning streak claiming they will keep winning forever. This approach appears particularly ill-considered in light of high-profile fatalities that already have occurred involving Uber and Tesla.

We should not be giving developers a free pass on safety until more people are injured or killed. This is especially true for testing practices that in effect use safety drivers as a moral crumple zone.[154]

MYTH #12: OTHER STATES/CITIES LET US TEST WITHOUT ANY RESTRICTIONS, SO YOU SHOULD TOO.

In the 2018 Tempe Arizona Uber ATG fatality a pedestrian was struck and killed by an AV test vehicle. Initial reports blaming pedestrian behavior and road lighting conditions were later discredited. While safety driver inattention contributed to the mishap, the root cause was unsafe testing practices that manifested as a symptom of a deficient safety culture.[155] The most recent version of SAE J3018 for road testing safety incorporates lessons learned from that tragic fatality. If testers won't follow that consensus industry standard, they haven't really taken those lessons to heart.

Regulators should pause to consider the consequences of putting vulnerable road users at increased risk to benefit for-profit companies. Those companies are using public roads as an experimental testing ground in their high-stakes race to autonomy. While road testing brings with it jobs and prestige for being tech-friendly, even a single testing fatality can draw worldwide negative attention to a region.

---

154. *See* Elish, *supra* note 127.

155. NAT'L TRANSPORT. SAFETY BD., ACCIDENT REPORT NTSB/HAR-19/03, COLLISION BETWEEN VEHICLE CONTROLLED BY DEVELOPMENTAL AUTOMATED DRIVING SYSTEM AND PEDESTRIAN 38 (2019), https://www.ntsb.gov/investigations/AccidentReports/Reports/HAR1903.pdf.

Regulators charged with ensuring safety should not feel inhibited from merely asking developers to follow the industry safety standards that in most cases the companies themselves helped write or had the opportunity to comment upon.[156] At-risk road users should not be used as unwitting test subjects for AV testers that, based on their actions, are not truly putting safety first. The combination of a failure to follow industry standards, coupled with the promotion of false and questionable narratives, is a practice that corrodes trust.

IV. "AUTONOMANDERING" AND THE CHALLENGE FOR THE POLITICAL PROCESS

This Part describes a practice in the AV industry which we call "autonomandering."[157] It bears a family resemblance to gerrymandering[158]—the portmanteau combining "autonomous" with "gerrymandering." Like gerrymandering, autonomandering ought to be disfavored because it poses challenges for a representative democracy. Engaging in the practice provides a further reason to withhold trust.

Autonomandering is a practice in which one or more AV companies lobby members of a state legislature to approve permissive statewide AV laws and regulations to preempt (and thus avoid) more restrictive, safety conscious municipal laws tailored for the special circumstances of an urban environment. It might be used proactively to block future more protective and thus restrictive municipal laws, as well as displace existing municipal

156. The On-Road Automated Driving (ORAD) SAE committee issued SAE J3016, SAE J3018, and other AV-relevant standards. The committee roster includes broad industry participation: https://www.sae.org/servlets/works/roster.do?comtID=TEVAVS (Registration with committee required for access).

157. We style the name of this practice "autonomandering" after the practice of "autonowashing" coined by Liza Dixon. *See supra* note 25. The term "autonowashing" describes the gap between the presentation of information about partial automated driving systems by the media and AV marketing teams and the actual system capabilities. This practice influences public perceptions of vehicle automation causing overreliance on partial automated driving systems, thus presenting a safety risk.

158. The harmful effects of gerrymandering are well known. "Gerrymandering refers to the drawing of political boundaries to favor one party, or one faction or another." Elaine Kamarck, *Gerrymandering and How to Fix It*, BROOKINGS (Feb. 2, 2018), https://www.brookings.edu/blog/unpacked/2018/02/02/gerrymandering-and-how-to-fix-it/.

regulation (such as the protective regulations found in New York City).[159] The strategy of passing a state law to neuter a local law is ubiquitous across subject matters, impacting areas of active public debate, such as minimum wage laws, gun regulations, status as a sanctuary city, and the ride hailing business.[160]

Particularly because the ride hailing business figures in the business plans of AV companies,[161] we predict that widespread autonomandering cannot be far behind. Indeed, since we first posted a version of this article on SSRN in November of 2021, legislators in Pennsylvania have proposed a law governing AV testing and deployment which has broad pre-emption provisions, which would effectively prevent Pittsburgh and Philadelphia from passing municipal laws appropriate for conditions in their urban environments.[162] The ride hailing business is only likely to be profitable (and certainly will be most profitable) in urban areas, so urban testing is a business necessity because it precedes deployment at scale in an urban environment. Accordingly, the AV industry would like to avoid any regulatory requirements which might impact or delay testing in urban environments.

This presents a challenge for the democratic process because the risk created by AV testing falls disproportionately on urban populations yet it prevents municipalities from responding to safety concerns expressed by

---

159. *See supra* note 15.

160. A website even tracks this practice. *Preemption Conflicts Between State and Local Governments*, BALLOTPEDIA, https://ballotpedia.org/Preemption_conflicts_between_state_and_local_governments (last visited Mar. 18, 2022). *See also* Spencer Wagner, *Three Strategies to Restore City Rights in an Era of Preemption*, NAT'L LEAGUE CITIES [NLC] (Dec. 13, 2019), https://www.nlc.org/article/2019/12/13/three-strategies-to-restore-city-rights-in-an-era-of-preemption/; SPENCER WAGNER, ET AL., NLC, RESTORING CITY RIGHTS IN AN ERA OF PREEMPTION (2019), https://www.nlc.org/wp-content/uploads/2019/11/Restoring-City-Rights-in-an-Era-of-PreemptionWeb.pdf (discussing tools and strategies to advance local decision-making in the face of preemption).

161. *See, e.g.*, Aurora S-1, *supra* note 1, at 83 (noting that Aurora plans to target the ride hailing business in 2024).

162. See S.B. 965, 2022 Gen. Assemb., Sess. 2022 § 8510 (Pa. 2022), https://www.legis.state.pa.us/cfdocs/billInfo/billInfo.cfm?sYear=2021&sInd=0&body=S&type=B&bn=965. Legislation has also been introduced in Kansas and Oklahoma.

their constituents—concerns which may not apply in other areas of a state.[163] IEEE 7000 states that "designers [of an AI system] need to take particular care that the system design and algorithms do not unjustly favor or select users in certain geographic areas . . . ."[164] Vulnerable residents of at least some urban areas did not implicitly consent to a higher risk exposure than rural residents, as evidenced by the passage of municipal laws more protective than general laws applicable statewide.[165]

Risk imposed upon other road users differs in kind from the risk exposure assumed by passengers who voluntarily enter an AV. A more protective municipal law, if enforced, might equalize the relative risk exposure between urban constituents and rural constituents (where more permissive testing might occur based on less restrictive laws applicable statewide). When AV companies engage in autonomandering, we consider it further evidence that the AV industry does not deserve public trust.[166]

Because this legislation has been proposed for Pennsylvania, its passage (let alone the lobbying effort to achieve pre-emption reflected by introduction of the bill) has the potential to taint the many AV companies located in Pittsburgh and elsewhere in Pennsylvania. With appropriate amendments, however, the bill might evolve into a model which helps establish Pennsylvania (or, with municipal regulation, Pittsburgh) as the epicenter for cutting edge technology developed the right way (for example, by following IEEE 7000 and allowing voices in urban areas to be heard).

In addition to IEEE 7000's concern that the development of AI systems avoid creation of geographic discrimination, we think it equally important

---

163. There are also potential equity issues because pedestrians in low-income urban areas are more vulnerable to death. *See* Tanya Snyder, *Study: People in Low-Income Areas More Likely to be Killed While Walking*, STREETSBLOG USA (Aug. 5, 2014), https://usa.streetsblog.org/2014/08/05/study-people-in-low-income-areas-more-likely-to-be-killed-while-walking/.

164. *See* IEEE 7000, *supra* note 5, at 27.

165. *See* text accompanying note 159 and *supra* note 15.

166. The video of the press conference appears on Senator Langerholc's webpage in which certain of the Senator's "legislative partners" are identified. https://www.senatorlangerholc.com/2022/01/05/langerholc-introduces-legislation-to-create-a-roadmap-for-highly-automated-vehicles/.

that the background conditions under which AI testing and deployment occur respect the democratic process, just as the operation of specific AI technology ought to do. The *Trustworthy AI Guidelines* emphasize the importance of democracy and the rule of law:

> A future where democracy, the rule of law and fundamental rights underpin AI systems and where such systems continuously improve and defend democratic culture will also enable an environment where innovation and responsible competitiveness can thrive.[167]

The concern expressed in the *Trustworthy AI Guidelines*—that AI development fosters democratic culture—raises the question of how the process of passing laws and regulations governing AV testing and deployment is proceeding throughout the United States.

The AV industry often focuses on interaction among federal, state, and local laws and regulations. The industry hope is that uniform standards might apply throughout the nation. (Achieving the goal is a step toward a uniform international standard.) In either case, uniform standards should foster both innovation and safety. From the AV industry perspective, any regulatory scheme should include a healthy dose of self-regulation.

The general concern expressed by the AV industry is that the United States presently has a patchwork of potentially inconsistent laws and regulations which might impede the rapid development of AV technology and hinder innovation.[168] Currently, the U.S. Congress has not passed any federal laws specifically regulating AV technology. However, existing federal laws and regulations might impact testing and deployment—for example, by requiring the presence of steering wheels[169] (which some AV

---

167. *Trustworthy AI Guidelines*, *supra* note 6, at 9. These guidelines were introduced, in part, to show "the right way to build a future with AI." *Id*.

168. *See, e.g.*, Michele Kyrouz, *Industry Comments to NHTSA's Federal Automated Vehicles Policy*, MEDIUM (Jan. 13, 2017), https://medium.com/smart-cars-a-podcast-about-autonomous-vehicles/industry-comments-to-nhtsas-federal-automated-vehicles-policy-436e7e24911a; Chris Giarratana, *Is the Battle of Autonomous Car Regulations Killing You?*, TRAFFIC SAFETY STORE (Dec. 1, 2016), https://www.trafficsafetystore.com/blog/autonomous-car-regulations/#.

169. *See generally* Removing Regulatory Barriers for Vehicles with Automated Driving Systems, 84 Fed. Reg. 24433 (proposed May 28, 2019), https://www.federalregister.gov/documents/2019/05/28/2019-11032/removing-regulatory-barriers-for-vehicles-with-automated-driving-systems.

manufacturers would like to eliminate from future products,[170] but which does not present insurmountable challenges to autonomy technology development efforts). Also, the FTC regulates advertising of AVs.[171] Moreover, federal regulators might proceed by granting exemptions from existing law and regulation to allow for limited public highway testing of certain products.

But as a general matter, the federal government has remained on the regulatory sidelines. A consensus is emerging that it will take years, if not a decade, before any meaningful and comprehensive federal legislation will be enacted, together with adoption of proper supporting interpretive rules and regulations approved after notice and comment. That is to say, the federal government is unlikely to take any proactive and forward-looking approach to AV safety. Rather, the prediction is that federal agencies will simply react to accidents by launching investigations and mandating data collection in the near and intermediate term.

Additionally, the current division of responsibility for safety is such that federal regulations cover automotive equipment, while state regulations cover the human drivers. Testing that involves human driver oversight places the vast majority of safety responsibility on the human driver, and thus should properly be in the realm of state regulation to the degree that mishaps are attributed to driver error rather than equipment failure.[172] When an automated driver replaces the human driver, this will change the balance of regulatory input from the states to the federal government if the automated driving system is treated as automotive equipment rather than akin to a human driver.

---

170. *See* Stephen Loveday, *Report: Elon Musk Says $25K Car Coming in 2023, Maybe No Steering Wheel*, INSIDEEVS (Sept. 3, 2021, 9:20 AM), https://insideevs.com/news/530786/elon-musk-tesla-compact-car-no-steering-wheel/.

171. William H. Widen, *Machine Driver Vs. Human Driver in Possible FTC Action Against Tesla*, JURIST (Aug. 26, 2021, 9:00 AM) https://www.jurist.org/commentary/2021/08/william-widen-machine-driver-vs-human-driver-ftc-tesla/.

172. The separation of driver responsibility from equipment failure becomes murky when considering whether driver monitoring system (DMS) technology is fit for its purpose. However, if the driver is blamed for a crash, then the basis for the blame clearly falls within the realm of state regulation.

Motivated, in part, by the absence of federal leadership, many states have passed laws regulating AV testing and deployment—some permissive (such as Arizona, Florida, Nevada and Texas),[173] and others with content intended to improve safety (such as California).[174] Recently, some local governments also have passed laws relating to AV testing, a signature example being New York City.[175] The primary local concern ought to focus on testing of AV technology (and less with deployment at scale).[176] In order to have a workable statewide and national transportation system, good reasons exist for uniform standards once AV technology is ready for deployment at scale.[177]

If, as in the case of New York City, a municipality passes a protective law imposing material requirements and conditions on the testing of AVs, an AV company might find such a law inconvenient. This is particularly true because the AV industry needs to engage in urban testing to capture the bulk of ride hailing business. How might an AV company remove a municipal law that makes testing inconvenient, or one that mandates publication of safety data (and prevent any future municipal regulation)?

The autonomandering strategy might proceed as follows. One or more AV companies draft a model statute (working together or through an industry group like PAVE) governing AV testing and deployment which contains few or no meaningful provisions supporting safety. For show, it contains some precatory language which, for marketing purposes, can be defended as safety conscious or safety aware. It might well be dressed up with an impressive recitation of technical definitions and a roll call of potential safety practices—but without requiring AV companies to conform to any of those safety practices.

The AV companies then present the model statute, first in private, to members of the target state legislature, hoping to drum up support for passage of a permissive bill. For illustrative purposes, assume the targeted

173. *See supra* note 13.

174. *See supra* Section I (discussing California law and regulations as applied to Tesla).

175. *See supra* note 15.

176. The issue is arguably more complicated when considering the question of whether an un-crewed test vehicle driving recklessly due to defective software would be a state or federal regulatory concern.

177. Nonetheless, such standards should preserve state and municipal prerogative to set local traffic regulations necessitated by local conditions as they do today.

members of the legislature represent more conservative and pro-business constituents who live in rural areas. The idea is to prearrange votes for the model statute. When the votes are lined up, the bill is introduced and placed on a fast track for enactment. If, through gerrymandering or otherwise, the targeted legislators command a majority in the state legislature, the model statute passes quickly, without meaningful hearings or debate.

The equitable problem with such an approach is that it promotes the interests and wishes of constituents for whom AV testing is a lower safety concern. Ex hypothesis, the residents of the urban area comprising the city which passed the municipal regulation have a greater concern over safety rather than potential economic concerns such as business development. It is often the case that citizens living in urban areas have a more progressive attitude toward the value of regulation than citizens living in rural areas.[178]

The net effect of autonomandering, should it occur, is that the state representatives of citizens at lower risk of harm defeat protective municipal regulation designed to address the specific safety concerns of the urban residents. This situation occurs if the passage of the model statute at the state level preempts more protective municipal legislation.

The easy fix for this inequitable situation is for the state legislation to specifically provide that it is passed in addition to, but not in lieu of, specific municipal regulation which addresses circumstances of local concern—such as higher traffic volume, greater road user vulnerability, equity concerns for road users put at risk, and more challenging situations in city driving conditions. It is particularly important that any statewide legislation expressly allow for local variation with respect to the testing of AV technology. This is the only way to equalize the risk exposure between urban residents and rural residents. By the time of deployment at scale, AV technology ought to be sufficiently developed so that urban residents do not face increased risks due to an immature technology, thus allowing for uniformity of treatment following deployment after testing.

---

178. *See, e.g.*, Timothy Callaghan, et al., *Rural and Urban Differences in COVID-19 Prevention Behaviors*, 37 J. RURAL HEALTH 287 (2021) (noting that rural residents are significantly less likely to participate in several COVID-19 related preventive health behaviors); ROBERT BONNIE, ET AL., NICHOLAS INST. FOR ENV'T POL'Y SOLS., DUKE UNIV., UNDERSTANDING RURAL ATTITUDES TOWARD THE ENVIRONMENT AND CONSERVATION IN AMERICA 17 (2020) (noting that "[r]ural respondents were much more likely to prefer less government oversight of environment and conservation issues compared to urban and suburban voters").

It is easy to withhold trust from companies that engage in autonomandering, even though this form of lobbying is commonplace. Given the many prominent AV companies located in Pennsylvania (such as Argo AI, Aurora and Locomation Inc.), the entire industry will present poorly if the proposed Pennsylvania bill goes forward without amendment—it is not a case of a single rogue actor behaving poorly, but the concerted effort of a group of prominent AV companies. What distinguishes AV regulation from other forms of legislation is the presence of ethical standards for AI development which demand that special respect be paid to democratic processes. Autonomandering fails to recognize or respect the ethical values represented by democracy.

V. PROBLEMS WITH J3016 AS A SAFETY STANDARD AND THE WAY FORWARD

This Part describes shortcomings of SAE J3016 in defining the scope of laws and regulations covering testing or deployment of AVs (whether applicable to a particular vehicle type or to a specific driving automation system or feature). The following discussion makes clear why use of J3016 as part of an AV industry regulatory scheme provides an additional, systemic reason not to trust that AVs will be tested and deployed safely. This is particularly true in a legal regime, as in the United States, where the default rule is that any action which is not prohibited is permitted.

This motivates our suggestion for a shift in approach to determining the scope of AV regulation towards a model which could better build systemic trust by creating more certainty of application. We suggest a modified version of SAE J3018 as a basis for AV testing regulation.[179]

A. Issues with J3016 to Define the Scope of a Law or Regulation

SAE J3016 sets forth a taxonomy for motor vehicle automation ranging from Level 0, representing no automation, to "full" automation at Level 5. Levels 0, 1, and 2 constitute "lower" levels of automation—respectively, no automation, driver assistance, and partial automation. Levels 3, 4, and 5 constitute "advanced" levels of automation—respectively, "conditional" automation, "high" automation, and "full" automation.[180]

As of this Article, the SAE has issued four versions of J3016 (in 2014, 2016, 2018 and 2021). Though the length of J3016 has increased from twelve

---

179. Regulation of testing is more problematic than regulation of deployment. The NHTSA DEC. 3, 2020 ANPRM, *supra* note 136, is a workable starting point for regulating deployed systems.

180. *See* J3016:2021, *supra* note 27.

pages to 41 pages, these revisions, while substantial (including clarifications, additional definitions, examples, and explanations), preserve the original SAE J3016:2014 level names, numbers, functional distinctions, and supporting terms.

Unfortunately, as discussed below, some of the revisions have reduced certainty over the scope of application of the different levels, rather than enhancing it. The SAE introduced these obfuscating revisions when it transitioned the purpose of J3016 from a mere taxonomy to facilitate technical discussions (as in J3016:2014) to a taxonomy that additionally might be used to set the scope of laws and regulations governing AV testing and deployment[181] (from J3016:2016 and forward).[182] Ironically, the original J3016:2014 might have provided a better basis for regulatory use by describing levels of driving automation systems and features in more objective terms because it does not expressly consider design intent.

1. Regulatory Boundaries: Vehicles or Driving Automation Features?

To date, regulatory efforts tend to rely on SAE J3016 to set a boundary between lower-level automation technologies that escape special AV regulation, and those higher levels of automation technology covered by AV-specific regulations. Regulators typically want to regulate the testing and deployment of *vehicles* whereas SAE J3016 technically is intended to classify levels of *individual features* provided by a driving automation system and not levels of vehicles. For example, the California law and

---

181. SAE went so far as to testify before US Congress in a bid to have the SAE Levels adopted as the basis for regulations. See Jennifer Shuttleworth, *Seeking a Common Language for Vehicle Automation*, SAE INT'L (May 24, 2017), http://articles.sae.org/15462/.

182. It appears to one of the authors with legal practice experience that the obfuscating revisions might have been introduced for the purpose of providing the AV industry with "flex" in the boundaries of the levels (making the determination of the boundary more subjective and less objective).

The AV industry might be more comfortable with allowing use of J3016 to define the scope of laws and regulations applicable to it with flex for a variety of reasons, including the ability to argue that any violation was unintentional. The additional flex gives any law or regulation which refers to a revised version of J3016 less force as a "safety guardrail" constraining the actions of industry participants. Of course, reducing the efficacy of a law or regulation as a safety guardrail reduces trust, rather than enhancing it.

regulations covering AVs focuses on testing and deployment of vehicles (rather than driving automation features).[183]

Vehicles deemed to be Levels 0, 1, or 2 escape special regulatory scrutiny because of their relatively low level of autonomous capability. Vehicles deemed to be Levels 3, 4, or 5 have higher autonomous capability, subjecting them to special regulation as AVs (at least in locations where state or local governments desire some form of enhanced oversight for high automation levels). An uncertain boundary between high and low levels of automated driving systems allows AV industry participants to explore, test, and expand the unregulated space for testing and deployment. Expansion of the unregulated space tends to reduce safety.[184]

Uncertainty regarding the scope of law and regulation provides a ready excuse for non-compliance by an industry participant because a participant might make a plausible claim that the uncertainty created involuntary non-compliance. As illustration, imposing criminal liability typically requires *scienter*, yet uncertainty surrounding the scope of application of a law or regulation makes proof of *scienter* more difficult. (Most business actors, whether natural or artificial persons, are more solicitous about criminal liability.)

Moreover, the mismatch between J3016's focus on levels of driving automation features and the regulatory agency's interest in regulating vehicles (rather than driving automation features) can cause confusion. To understand how this confusion might develop, consider the complexity of the J3016 structure itself. When defining the scope of a law or regulation, complexity is undesirable, particularly if certainty is the goal.

---

183. For our purposes we informally define the "level" of a *vehicle* to be the level associated with the highest-level feature the vehicle is capable of, even if that feature is only activated a small fraction of the time the vehicle is being operated. This definition seems consistent with regulatory approaches.

184. Though expansion of unregulated space tends to decrease safety, the AV industry frequently argues that an expanded unregulated space is necessary for innovation and faster deployment of valuable technology. The question of whether a boundary is clear as a legal and administrative matter is separate and distinct from the question of where that boundary ought to be set as a matter of policy. One could reasonably expect that unclear boundaries are corrosive to regulatory authority and effectiveness, regardless of where any boundary might be set.

Under J3016, a vehicle might employ a driving automation system or feature at Level 4, intended for use in a limited operational design domain (ODD). For example, a driving automation system might be designed for use only on interstate highways in dry weather. As a Level 4 feature, in the language of J3016, this would be an Automated Driving System (ADS)[185] and not merely a driving automation system because it can perform the complete dynamic driving task (DDT) on a sustained basis.[186]

Such a feature might contemplate that a driver operates the vehicle manually, without use of automation, from her house to the on-ramp of the interstate (segment 1; operated at Level 0), engaging the Level 4 driving automation system on the interstate (segment 2; operated at Level 4), and reverting to manual operation to exit the interstate, next continuing on local streets to the destination (segment 3; operated at Level 0). During the Level 4 segment of the trip, the human driver might read a book or take a nap, leaving the entire DDT to the driving automation system (making this system or feature an ADS even though used for only a portion of the itinerary).[187]

Testing of such a vehicle ought to be subject to safety regulation as a Level 4 *vehicle*, even though use of that driving automation system is appropriate for only part of a trip, because the vehicle is, nevertheless, equipped with an operational ADS feature.[188] The risk to others using the interstate is the same during segment 2 of the trip regardless of the feature's operational status during segments 1 and 3. The various versions of J3016 specifically contemplate that different trip segments might engage different levels of driving automation systems such as in this interstate ODD

---

185. The abbreviation "ADS" stands for "Automated Driving System and applies to Levels 3, 4 and 5 only. Per J3016, an ADS must support features at SAE Levels 3, 4 or 5. J3016:2021, *supra* note 27. The term "driving automation system" is a more generic term encompassing lower SAE Levels as well. *Id.* at 6. The lower-level systems perform *part of* the Dynamic Driving Task (DDT) on a sustained basis whereas the higher-level systems perform *all* the DDT on a sustained basis. *Id.* at 6–7. Only a vehicle limited to Level 0 features is without a driving automation system. *Id.* at 4. In J3016:2014, the term "automated driving system" was sometimes used generically before the definition of ADS was introduced in 2016. J3016:2014, *supra* note 27.

186. *See* J3016:2021, *supra* note 27, at 20, 26 tbl.1, 40.

187. *See id.* at 8 fig.1.

188. *Id.* at 8.

example.[189] However, the public discourse often fails to note that a vehicle can engage a high level of automation for only a portion of a trip, yet remain subject to regulation as an AV by virtue of operation at Level 3 or above for only part of an itinerary. But it is not true that the inability of a vehicle to always operate at a high level of automation, in all conditions, renders the vehicle Level 2.

Only a vehicle equipped with Level 5 driving automation features needs to operate across all ODDs. Even then, an "ADS feature designed by its manufacturer to be Level 5 would not automatically be demoted to Level 4 simply by virtue of encountering a particular road on which it is unable to operate the vehicle."[190] Under revised versions of *J3016*, the level of a driving automation system or feature corresponds to the manufacturer's design intent for the production version of the feature.[191] Design intent may exceed actual "real world" performance.[192]

In the interest of clarity, if a law or regulation purports to regulate *vehicles* rather than driving automation systems (particularly those that also are ADS), the law or regulation should also refer to a descriptive schema focused on vehicles rather than features. Alternately, the law or regulation should make clear that a vehicle which can operate during any segment of

---

189. *See, e.g.*, *id.* § 5.5 n.2, at 32.

190. *Id.* § 8.2, at 36.

191. *Id*. J3016:2021 states most clearly the importance of design intent to the determination of a level. However, prior versions of J3016 have relied on the concept of design intent to set a level as well. *See* J3016:2016, *supra* note 27, at 27 (noting that "[t]he level assignment rather expresses the design intention for the feature") (emphasis omitted); J3016:2018, *supra* note 27, at 30 (stating that "[l]evels are assigned, rather than measured, and reflect the design intent for the *driving automation system feature* as defined by its manufacturer"). The emphasis on design intent to determine level does not appear in J3016:2014.

192. *See, e.g.*, J3016:2016, *supra* note 27, at 2 (noting that "'Role' in this context refers to the expected role of a given primary actor, based on the design of the driving automation system in question and not necessarily to the actual performance of a given primary actor.") (emphasis omitted). "Driving" involves three primary actors: the human driver, the driving automation system and other vehicle systems and components. It is worth noting that a dangerous AV that drives recklessly might be classified as a Level 4 vehicle based on design intent, even if a human safety driver is kept constantly busy intervening to avoid crashes in that vehicle's current technological incarnation.

a trip using an ADS should be regulated as an "autonomous vehicle," as defined in the law or regulation.

2. J3016 is Not a Safety Standard But "Something Else"

The SAE has promulgated four versions of J3016: in 2014, 2016, 2018 and 2021. No version purports to be a safety standard *per se* (in contrast with SAE J3018[193] for example). The original J3016:2014 was a modest document of twelve pages, denominated as an "Information Report." An information report is neither an SAE recommended practice nor a standard.[194] The original J3016 explained that it was published to "provide[] a foundation for further standards development activities and a common language for discussions within the broader 'Automated/Autonomous Vehicle' community."[195] Importantly, J3016:2014 expressly disclaimed use for legal purposes: "SAE's levels of *driving* automation are descriptive rather than normative and technical rather than legal."[196] It also clarified that its scope was limited to "provid[ing] a taxonomy describing the full range of levels of automation in on-road *motor vehicles*. It also includes operational definitions for advanced levels of automation and related terms."[197]

With the first revision (in 2016), J3016 increased to 30 pages in length, and its status was upgraded from "Information Report" to "Recommended Practice."[198] J3016:2016 included two principal types of additions. First, it included more fulsome operational definitions for the lower levels of autonomy, which had been omitted from J3016:2014. Second, it added clarifications through more definitions and examples. With these additions, the SAE expanded the purpose of J3016:2016, indicating that it now had potential legal application: "Standardizing levels of driving automation and supporting terms serves several purposes, including . . . [a]nswering questions of scope when it comes to developing laws, policies, regulations,

---

193. *See* J3018, *supra* note 28.

194. *See Standards Development Process*, SAE INT'L, https://www.sae.org/standardsdev/devprocess.htm (last visited Mar. 18, 2022).

195. J3016:2014, *supra* note 27, at 1.

196. *Id.* § 3, at 2.

197. *Id.* at 1.

198. J3016:2016, *supra* note 27, at 1.

and standards."[199] Further, J3016:2016 indicated that its revisions should "[b]e useful across disciplines, including engineering, law, media, public discourse."[200] This same purpose and scope appears in the 2018 and 2021 versions of J3016.[201]

Despite the expansion of its purpose to answer questions of scope for developing laws and regulations, the basic structure of J3016 did not change. It remains today "descriptive and informative, rather than normative, and technical rather than legal,"[202] just as it was in 2014.[203] But the revisions made J3016 less suitable to define the scope of a law or regulation with any certainty by, among other things, introducing the concept of "design intent" as one factor in setting a level.[204]

The expansion of J3016's purpose to include potential legal application was done by the mere stroke of a pen without justification based on a change in its structure or approach. Indeed, for two principal reasons, J3016 is deficient for the purpose of setting the scope of a law or regulation. First, the revisions to J3016 rely on the subjective notion of "design intent" to give content to the various automation levels.[205] Yet determining the scope of a law or regulation with legal certainty requires reference to objective, rather than subjective, criteria.

Use of an intent standard introduces familiar evidentiary challenges. It requires development of objective signs to prove subjective facts if it is to have practical legal force. This is obvious in the case of the intent of a natural person because we cannot directly inspect the contents of another mind.

---

199 *Id.* (emphasis omitted).

200. *Id.*

201. J3016:2018, *supra* note 27, at 1; J3016:2021, *supra* note 27, at 1.

202. J3016:2021, *supra* note 27, § 4, at 24. The same limitation appears in J3016:2018, *supra* note 27, § 4, at 18 and in J3016:2016, *supra* note 27, § 4, at 17.

203. J3016:2014, *supra* note 27, § 3, at 2.

204. J3016:2014 uses the term "design intent" in two places, neither of which impact the setting of a level. *Id*. at 10.

205. The term "design intent" is not defined within the standard, and no external reference is provided to assist in defining the term.

The situation is further complicated, however, because at issue is the intent of a corporation or other business entity in the AV industry.

Artificial entities only operate through human agents such as employees, managers, and directors. Does one determine the intent of an artificial entity by reference to the intent of the design engineers, the marketing department, upper-level management, or the board of directors (or some combination which must be weighed and balanced in some yet to be determined fashion)? And, of course, different natural persons within the same organization might have different design intents, each of which might have a claim for attribution to the organization.

Moreover, a law or regulation which simply takes an assertion about design intent at face value, without some independent confirmatory sign, is very weak indeed. As a comparison, in contract law the meaning of a contractual term is determined by reference to objective criteria related to external signs and usages from the perspective of a reasonable person (rather than testimony by a party about his or her subjective intent).[206] Any other approach would make contracts illusory and render the institution of contracting of little value for business planning.

Second, J3016 now looks first and foremost to the manufacturer (not the regulator) to assign the automation level—and thus, indirectly, the scope—to a particular driving automation system. When combined with use of design intent to establish a level, AV companies can, effectively, design their own regulatory regime. As a general matter, however, interpretation of a law or regulation is a matter for a court or a regulator, not a private party.

The judgements of the various AV industry participants might differ in their respective determinations of scope. AV Company No. 1 might classify a driving autonomy system as Level 4 (based on its "design intent"—even if the current, immature incarnation requires constant driver supervision for safety), while AV Company No. 2 might classify a feature with substantively identical capabilities as Level 2. Allowing multiple private parties to set potentially inconsistent standards for the same technology feature creates confusion, making uniform safety regulation a virtual impossibility. Moreover, such a system is subject to manipulation by

206. *See, e.g.*, Wayne Barnes, *The Objective Theory of Contracts*, 76 U. CIN. L. REV. 1119, 1119–20 (2008) (noting that the objective theory of contracts provides that mutual assent to a contract is determined by reference to external acts and manifestations, not by evidence of subjective, internal intention).

industry participants, particularly if there is a business advantage to be gained by doing so.

When a law or regulation relies on a determination made by a private party (far from the norm), disaster often follows—for example, the financial meltdown in 2008 which was due, in no small part, to the incorporation into law and regulation of securities ratings provided by private rating agencies.[207] The rating agencies' judgment had been compromised by financial pressures, which resulted in inappropriately high ratings for many securities, creating increased risks for investors. Financial pressures similarly might influence the decision of an AV company (whether consciously or unconsciously) to set the level of a driving automation system at Level 2 to evade regulatory oversight.

Nevertheless, following the SAE's invitation to use the revised 2016 version of J3016 to answer questions of scope for laws and regulations, the California DOT incorporated J3016:2016 into regulations governing testing and deployment of AVs. And the California legislature recently followed this approach by incorporating J3016:2021 into the related statute.[208] To be sure, there is nothing wrong with incorporating technical materials into regulations. This practice has many benefits. Indeed, it is a preferred method used by the U.S. government in the Code of Federal Regulations to efficiently promulgate regulations.[209] However, incorporation by reference

---

207. *See* Randall D. Guynn, *The Financial Panic of 2008 and Financial Regulatory Reform*, HARV. L. SCH. F. ON CORP. GOVERNANCE (NOV. 20, 2010), https://corpgov.law.harvard.edu/2010/11/20/the-financial-panic-of-2008-and-financial-regulatory-reform/.

208. CAL. CODE REGS. tit. 13, § 227.02 (2021); *accord* CAL. CODE REGS. tit. 13, § 228.02(b) (2021) (stating that the definition of "autonomous vehicle" meets SAE Levels 3, 4, or 5); 2021 Cal. Legis. Serv. Ch. 277 (S.B. 500) (West) (2021 portion of 2021-2022 Regular Session updating § 38750 as of Sept. 23, 2021).

209. Congress authorized incorporation by reference in the Freedom of Information Act to reduce the volume of material published in the Federal Register and CFR. *See* 5 U.S.C. § 552(a); and Incorporation by Reference, 1 C.F.R. pt. 51. Incorporation by reference raises many issues which have been discussed elsewhere in the literature. *See, e.g.*, Emily S. Bremer, *Technical Standards Meet Administrative Law: A Teaching Guide on Incorporation by Reference*, 71 ADMIN. L. REV. 315 (2019). These issues include whether, and to what extent, a regulation which refers to a specific version of a technical standard is automatically updated when the technical standard is updated. Automatic updating raises problems because the public did not have notice, or a chance to comment upon, the revisions in the

only succeeds to the extent the materials incorporated prove adequate for the task. The problem with J3016 is that it has proved neither adequate for legal purposes nor adequate for engineering, media, and public discourse.

3. The Apparent Failure of J3016 to Provide a Complete Vocabulary

J3016 has proved inadequate for at least some engineering, media, and public discourse, as evidenced by introduction of the term "Level 2 plus,"[210] used by some in AV discourse to characterize various advanced (but allegedly "low level") automation technologies, including Tesla's FSD. As noted above, the original J3016:2014 was intended to create "a common language for discussions within the broader 'Automated/Autonomous Vehicle' community."[211] The introduction of a seemingly technical term like "Level 2 plus" into the discourse (if truly needed) reflects a deficiency within the expressive power of the intended common language. Indeed, the two most recent versions of J3016 specifically state that it is incorrect to use qualified or fractional references to a level, such as 2.5 or 4.7: "Qualified or fractional levels would render the meaning of the levels ambiguous by removing the clarity otherwise provided by the strict apportionment of roles between the user and the driving automation system in performance of the DDT and fallback for a given vehicle."[212]

We infer that motivation to introduce the term "Level 2 plus" into the AV discourse comes from an industry desire to identify advanced automation technology for discussion purposes without leading to regulatory oversight that would follow from a feature being classified

---

newer version of the standard. As the California regulation referred to the 2016 version of J3016, the presumption is against automatic updating. In contrast, the amendment to § 38750 specifically purports to include updated versions of J3016 after September 23, 2021.

210. *See, e.g.*, Dean Takahashi, *Nvidia Launches Drive AutoPilot with Xavier AI Processors for Commercial Use*, VENTUREBEAT (Jan. 7, 2019, 2:00 PM), https://venturebeat.com/2019/01/07/nvidia-launches-drive-autopilot-with-xavier-ai-processors-for-commercial-use/ (noting that "Nvidia announced that its Drive AutoPilot is the first automated driving system that meets standards for Level 2-plus autonomous cars"). SAE International itself has even published articles using the designation "Level 2+." *See* Lindsay Brooke, *'Level 2+': Making Automated Driving Profitable, Mainstream*, SAE INT'L (Dec. 24, 2020), https://www.sae.org/news/2020/12/rise-of-sae-level-2.

211. *See supra* text accompanying note 189.

212. *See* J3016:2018, *supra* note 27, at 30 (emphasis omitted); *accord* J3016:2021, *supra* note 27, at 37.

above Level 2 in the SAE taxonomy. The ability to potentially sell advanced automation technology while still holding the driver responsible for any crashes likely plays a part as well.[213]

One industry participant indicated that the idea behind "Level 2 plus" is to describe an advanced, yet simultaneously low level, automation technology. (One might wonder whether conjoining the concepts of "advanced" with "low level" even makes any sense.) The purported basis for this added distinction is that "Level 2 plus" described "new systems that add safety and comfort features but always keep the driver in the loop."[214] Keeping the driver in the loop, however, has nothing to do with assignment of level. And the requirement of a safety driver does not, in and of itself, result in Level 2 status. The irrelevance of a driver in the loop appears in J3016:2018 (the applicable SAE document at the time of the errant remarks):

> As such, it is incorrect to classify a level 4 design-intended *ADS feature* equipped on a test *vehicle* as level 2 simply because on-road testing requires a test *driver* to *supervise* the *feature* while engaged, and to intervene if necessary to maintain safe operation.[215]

The SAE taxonomy does not make these fine gradations within a level. Use of a term like "Level 3 minus," for example, would not offer the same cosmetic effect to avoid the scope of regulation. What makes a technology Level 2 plus rather than Level 3 minus? If determination of level is based on the respective roles of human driver and driving automation system, intermediate or fractional gradations should not be necessary.

Use of the term "Level 2 plus" is perhaps intended to describe an unregulated zone between Level 2 and Level 3. A concrete example of an otherwise reasonable use of the term "Level 2 plus" might be to describe a Level 2 feature that includes both ODD enforcement (that the feature will engage only within its intended ODD) and an effective driver monitoring system—neither of which are required to qualify as a Level 2 feature, but both of which would substantively contribute to safety. Yet the taxonomy's

---

213. For a description of the practice of holding the driver responsible for crashes as a kind of "moral crumple zone" *see* Elish, *supra* note 127.

214. *See* Brooke, *supra* note 210.

215. J3016:2018, *supra* note 27, § 8.2, at 30. J3016:2021 contains the same observation in its version of § 8.2. J3016:2021, *supra* note 27, at 36.

stated goal is to avoid gaps of this sort.[216] Thus, there simply is no place for an additional "Level 2 plus" in the current J3016 scheme.

An additional concern with J3016 is that it excludes from its scope crucial aspects of practical safety for automated vehicles. The emphasis on J3016 is driving the vehicle; it excludes strategic aspects of vehicle operation such as route selection.[217] Further, J3016 does not address other critical aspects of vehicle safety for which a human driver would assume responsibility, but which go beyond the actual task of controlling vehicle motion. These include: post-crash scene responsibilities, ensuring under-age passenger safety, ensuring cargo safety, and ensuring the integrity of consumables (for example, tire tread depth, windscreen condition, vehicle lighting system operability), which are all important safety considerations that must be dealt with even in a vehicle that has no human driver.[218]

Additionally, J3016 has proved inadequate for legal purposes because of a mix of potentially bright line rules with seemingly flexible standards, requiring the exercise of judgment for application.[219] J3016 generally assigns levels based on a manufacturer's design intent (a non-obvious fact rather than an objective measure) and appears to cede assignment of level to the manufacturer itself.[220] When incorporated by reference into a law or regulation, however, reference to a subjective measure (such as intent) or a

216. For example, J3016:2016 states that the levels, definitions and terms "can be used to describe the full range of *driving automation features* equipped on *motor vehicles* in a functionally consistent and coherent manner," indicating by use of the word "full" that J3016's taxonomy was intended to be complete, without gaps. J3016:2016, *supra* note 27, at 2 (emphasis supplied).

217. J3016:2018, *supra* note 27, § 8.11, at 34.

218. The ANSI/UL 4600 system safety standard covers these types of considerations, especially in § 14 on lifecycle concerns.

219. For example, the subjective judgement of individual manufacturers as to assignment of levels creates potential for problematic inconsistencies. For a description of the difference between rules and standards, see Pierre Schlag, *Rules and Standards*, 33 UCLA L. REV. 379 (1985). The law governing the standard of conduct for a driver who comes to an unguarded railroad crossing illustrates the difference. *Id.* at 379. Oliver Wendell Holmes suggested a bright line rule: The driver must stop and look. *Id.* Benjamin Cardozo offered a flexible standard requiring the exercise of judgment: The driver must act with reasonable caution. *Id*.

220. J3016:2021, *supra* note 27, §5, at 30.

flexible standard (allowing the manufacturer to determine level which may vary by company) proves extremely problematic as a systemic matter because they impede legal certainty. The interpretation of the meaning and scope of J3016 ought to be transferred to the realm of the courts and the regulators following incorporation by reference. Typically, a regulator cannot delegate its responsibility of regulatory oversight to a private party.

Moreover, the tests for satisfying an autonomy capability level are specified by a mix of textual verbal definitions, summary charts, and examples, which include apparent gaps and inconsistencies.[221] The net result is that, in application, J3016 in its various iterations is both subject to manipulation by AV industry participants and confusing to lawmakers, regulators, and the public.

Shortcomings in J3016's structure in the original J3016:2014 has been exacerbated by subsequent revisions. However, a problematic organizational structure did not matter for J3016:2014 because the original version expressly disclaimed any legal application. This changed with J3016:2016 and beyond (with each subsequent version claiming applicability for legal use). The intended use changed but the revisions introduced problematic standards applicable to setting automation levels—the subsequent versions increased in length and added additional subjective measures, creating more uncertainty.

B. An Illustration from the History of J3016

Examining the history of the specification of low levels of automation in the different versions of J3016 illustrates the attempt to provide certainty—showing where it succeeded and where it failed. J3016:2014 contains potential bright line tests for Levels 0, 1, and 2. Despite the assertion that J3016:2014 contains functional definitions for only high automation

221. As examples, the specification of minimum standards to qualify for a level is not complimented by a specification of maximum capabilities that an ADS may have while remaining at a given level. It is unclear when a transient change in an operating environment represents an "ODD exit." J3016:2021, *supra* note 27, at 17. The line between sustained performance of part or all of the DDT and something other than sustained performance is unclear. And systems that provide momentary intervention in lateral and/or longitudinal motion control but do not perform any part of the DDT on a sustained basis (such as electronic stability control) are not classifiable (other than at Level 0) under the taxonomy. *Id.* at 20.

levels,[222] it in fact outlines functional definitions for low automation levels as well. These functional definitions receive clarification in later versions.

In J3016:2014, Table 2 describes the ability of the vehicle for Level 0 (No Automation) as follows: "No active automation (but may provide warnings)."[223] Table 2 describes Level 1 (Driver Assistance) as allowing the vehicle to execute portions of the dynamic driving task but limited to control of either longitudinal (accelerating, braking) or lateral (steering) motion, not both.[224] In Level 2 (Partial Automation), the vehicle executes both longitudinal and lateral aspects of the dynamic driving task when activated.[225] In both Level 1 and Level 2, the driver assistance/partial automation deactivates immediately upon request by the human driver.

A key difference between low automation (Levels 0, 1, and 2) and high automation (Levels 3, 4, and 5) is that in low automation, the human driver monitors the driving environment, whereas in high automation, the automated driving system monitors the driving environment when activated.[226] Even within J3016:2014 considered in isolation (before introduction of "design intent" to define levels) there are a few ambiguities of scope. Table 1 indicates that a Level 0 vehicle might "intervene" in addition to providing a mere warning, whereas Table 2 characterized Level 0 as merely providing warnings.[227] While Table 1 specifies that in Levels 1 and 2 the human driver "monitors" the driving environment, it also indicates that the vehicle's driving automation system[228] uses "information" about the driving environment—suggesting that a Level 1 or 2 vehicle monitors some aspects of the driving environment, which the human driver

---

222. J3016:2014, *supra* note 27, at 1.

223. *Id*. at 3 tbl. 2.

224. *Id*.

225. *Id*.

226. *Id*. at 3-4 tbl 2.

227. *Id.* at 2 tbl. 1; *Id.* at 3 tbl.2. J3016:2014 § 7.1 makes clear that "certain automatic emergency intervention systems" might be included and still allow a vehicle to be low automation*. Id*. at 9.

228. A "driving automation system" refers to any level 1-5 system or feature that performs all or part of the DDT on a sustained basis. *Id*. at 3-4 tbl 2. In contrast, the term "automated driving system" or "ADS" applies only to levels 3-5. *Id*. at 2 tbl. 1.

also monitors. J3016:2014 defines "monitor" as consisting of "activities and/or automated routines that accomplish comprehensive object and event detection, recognition, classification, and response preparation, as needed to competently perform the *dynamic driving task*."[229]

This definition makes clear that a low automation system is not engaged in monitoring (as defined) when it detects and processes limited information about the driving environment to control lateral vehicle motion and longitudinal vehicle motion, because monitoring requires a comprehensive assessment of the driving environment. Significantly, J3016:2014 uses the concepts of lateral and longitudinal vehicle motion only to make distinctions among low automation levels, not high automation levels.[230]

What changed with the revisions in J3016:2016? One purpose for the revisions was to clarify and rationalize the "taxonomical differentiator(s) for lower levels (levels 0-2)."[231] This required, among other things, an enhanced explanation of the type of information a low driving automation feature collects and processes for lateral and longitudinal motion and was achieved by indicating a standard that falls short of comprehensive monitoring of the driving environment.

Thus, J3016:2016 includes definitions for lateral and longitudinal motion: "*Lateral vehicle motion control* includes the detection of the *vehicle* positioning relative to lane boundaries and application of steering and/or differential braking inputs to maintain appropriate lateral positioning."[232] "*Longitudinal vehicle motion control* includes maintaining set speed as well as detecting a preceding vehicle in the path of the subject vehicle, maintaining an appropriate gap to the preceding vehicle and applying propulsion or braking inputs to cause the vehicle to maintain that speed or gap."[233]

Notice that J3016:2016 limits the scope of lateral and longitudinal vehicle motion control with respect to the elements of the driving environment

---

229. *See* J3016:2014, *supra* note 27, § 4.7, at 6.

230. *Id.* § 5, at 7.

231. J3016:2016, *supra* note 27, at 2.

232. *Id.* § 3.10, at 8.

233. *Id.* § 3.11, at 8 (certain emphasis omitted).

needed to execute the task. Lateral control consists of identifying lane boundaries to keep the vehicle in its lane. Longitudinal control consists of detecting a preceding vehicle in the path of the subject vehicle and maintaining an appropriate gap for a set speed. As defined, it does not explicitly include other aspects of the dynamic driving task, such as detecting stop signs and avoiding pedestrians, nor does it encompass detecting and avoiding collisions with cross traffic at intersections.[234] Further, even the concept of executing turns has been removed from later versions of J3016.[235]

Moreover, the combination of lateral and longitudinal control in a Level 2 driving automation feature does not appear to cover the execution of a lane change.[236] J3016:2016 captures this limited role by expanding the description of the dynamic driving task and introducing the abbreviation "OEDR" to refer to object and event detection and response.[237] A Level 2 vehicle only performs "limited OEDR associated with vehicle motion control."[238] Levels 3, 4, and 5, on the other hand, include "complete OEDR" because they perform the complete dynamic driving task when engaged in the applicable operational design domain.[239]

J3016:2016 also clarified the scope of intervention permitted for a vehicle to remain classified as Level 0.

> *Active safety systems*, such as electronic stability control and automated emergency braking, and certain types of driver

---

234. An ambiguity arises because the elements are intended to specify a minimum for a level, and not a maximum. *See infra* text accompanying note 232. For example, a vehicle that can handle 99.99% of the OEDR requirements with a driver completing the last 0.01% is still a Level 2 system, the same as a vehicle that is only capable of keeping itself centered in a highway lane at constant speed while only being able to follow a leading vehicle.

235. Indeed, the word "intersection" and even the concept of performing a turn does not appear in J3016:2016 or later versions at all, even though "turning" was within scope of the DDT in J3016:2014 § 4.4. *See* J3016:2014, *supra* note 27, § 4.4.

236. J3016:2014 made a distinction in its definition of Dynamic Driving Task between "lane keeping" and "lane changing." *See id.* at 6.

237. J3016:2016, *supra* note 27, at 5.

238. *Id.* at 6.

239. *Id*.

> assistance systems, such as lane keeping assistance, are excluded from the scope of this *driving automation* taxonomy because they do not perform part or all of the *DDT* on a *sustained* basis and, rather, merely provide momentary intervention during potentially hazardous situations. Due to the momentary nature of the actions of *active safety systems*, their intervention does not change or eliminate the role of the *driver* in performing part or all of the *DDT*, and thus are not considered to be *driving automation*.[240]

Thus, momentary control of lateral or longitudinal motion in an emergency (as opposed to control on a sustained basis) is merely an "active safety system" that is not considered to be "driving automation," even though there may be an element of "active control of a vehicle subsystem (such as brakes, throttle, suspension, etc.)."[241] Examples of momentary control include anti-lock brake systems, electronic stability control and automated emergency braking.[242] An exception to the requirement of momentary control is conventional cruise control, which is not momentary. However, conventional cruise control only maintains vehicle speed based on the vehicle's internal instrumentation rather than responding to external events in the driving environment, and thus, is classified as Level 0.[243]

The foregoing analysis suggests that Levels 0, 1, and 2 might be specified by objective standards and measures without reference to design intent or determination of level by a manufacturer. For example, Level 0 features provide warnings, not control of vehicle motion (with a limited exception for temporary motion control in an emergency). Levels 1 and 2 depend on whether an automated driving feature controls both lateral and longitudinal motion (or only one), using limited environmental inputs such as identification of highway lines and immediately preceding vehicles. Evaluation of these limited inputs is not monitoring the complete driving environment, but only a portion of it.

Conventional cruise control is not an automation feature because it does not respond to changes in the driving environment. Active safety systems

240. *Id.* at 2.

241. *Id*. at 3.

242. *Id*. at 13.

243. *Id*.

such as electronic stability control and automated emergency braking do not constitute a driving automation feature because they do not perform part or all the DDT on a sustained basis.

One might simply and objectively define the scope of a law or regulation by stating that any driving automation system or features in addition to the above are subject to regulation. The problem arises because under J3016 these "[e]lements indicate minimum rather than maximum capabilities for each level."[244] The subjective and uncertain elements of design intent and manufacturer specification creep into J3016 because J3016 fails to specify a maximum capability for each level, leaving that determination outside the elements specified in the four corners of the document. Without a specification of the maximum degree of autonomy capability that a Level 2 driving automation system might possess while remaining Level 2, the scope of application of a law or regulation will remain uncertain.

A driving automation system might attempt technical compliance to remain Level 2 by simply omitting some small aspect of the Level 3 specification by failing to monitor an obscure aspect of the driving environment. For example, an otherwise Level 3 system that cannot respond properly to an elephant walking on city streets, but intentionally considers circus parades within its ODD, might be claimed to be Level 2 because the "design intent" is to intentionally not detect elephants, rendering its OEDR response incomplete and therefore Level 2. This could be true even if it is solely operating in a geographical area in which a circus parade from the train station to a local tenting site is purely a theoretical event which has never actually happened and probably never will. This same technique might be used to maintain a lower level in other scenarios.

Lastly, the way that SAE has revised J3016 creates additional problems for use of J3016 across time. The SAE has denoted each revision with the letter "R" before the document name. Use of the "R" designation indicates that the document is a "complete" revision. A complete revision does not need to indicate substantive changes by the use of bars ("|") in the margin. However, despite the "R" designation, none of the revisions is a complete revision but rather a refinement or clarification. The failure to properly and completely identify those changes intended as a change of substance (as opposed to a clarification) makes use of the current publication take place

244. J3016:2021, *supra* note 27, at 24.

in the absence of any meaningful "legislative history." This circumstance leads to further confusion.

For all the foregoing reasons, J3016 falls short for use in specifying the scope of laws or regulations. The very structure of J3016 makes a regulatory scheme which incorporates it by reference uncertain in scope and untrustworthy in application. To get a better fit using objective rather than subjective factors, the focus on specification of scope should shift to an actual safety standard like J3018, with appropriate adjustment to delete the problematic reference to J3016. We now turn our focus to that exercise.

C. The Way Forward

1. Differentiating Testing from Production

As a practical matter, a significant problem with the application of J3016 to regulation is the conflation of AV testing with end-user operations. This allows a party to game the levels by manipulating declared "design intent" to suit marketing or regulatory goals.

Especially problematic is conflation of a Level 2 production vehicle with a Level 3, 4 or 5 human-supervised developmental test platform. These are superficially similar in that both have a human driver who is responsible for intervening if the system is unable to drive properly.[245] However, at a more nuanced level, the safety implications differ dramatically. An issue we call the "Level 2 loophole" is exploited when a manufacturer operating or selling access to what are really Level 3-5 test vehicles claims that they are only Level 2 vehicles, evading regulatory oversight applied to Level 3-5 technology.

We address the situation by categorizing technology for regulatory purposes independent from the SAE Levels. But to get there, we first cover differentiations that should be made between production Level 2 vehicles and higher-level test platforms, since the SAE Levels are the starting point for current regulatory approaches.

245. Human supervision is required in a Level 2 vehicle because it does not perform the entire dynamic driving task. Human supervision is used in a test vehicle to provide an extra layer of safety. *See* J3018.

2. Production Level 2 Regulation

With a production Level 2 system, a driver is not expected to control vehicle motion, but must watch for objects and situations the vehicle is not able to handle properly.[246] The driver intervenes if the vehicle is not able to detect an object or event, is unable to mount a safe response, or to recover the situation in case of vehicle equipment failure.[247]

Because J3016 is not a safety standard, it does not elaborate on the implications of the Level 2 approach regarding safety. More aggressive interpretations of J3016 treat any vehicle in which a driver is put in a driver seat as Level 2, which can lead to unreasonable risk in practice. To address safety, we propose that the following attributes be used to characterize a production Level 2 system for regulatory purposes:

1. Vehicle automation is capable of both lateral and longitudinal control, but not capable of executing turns at intersections,[248]
2. Any licensed driver should be able to operate a vehicle with net combined vehicle-plus-driver safety at least as good as for Level 0

---

246. *See supra* note 239 (differentiating function of a human driver in a Level 2 vehicle compared with a higher-level vehicle).

247. More technically, the human driver completes the Object and Event Detection and Response (OEDR) sub-task of driving, additionally performing in the role of Fallback driver if an equipment failure or exit from the ODD occurs.

248. This amounts to characterizing Level 2 as a super-smart cruise control that can perform both speed control and lane keeping, stopping for objects it is able to detect. Automated control of both lateral (steering) and longitudinal (speed) vehicle motion portions of the DDT is the basic definition of Level 2 automation. J3016:2016, *supra* note 27, tbl. 1. J3016:2014 § 4.4. included "steering, turning, lane keeping, and lane changing" as part of the DDT. J3016:2014, *supra* note 27, § 4.4. However, J3016:2016 omits "turning" from the DDT description and does not refer to vehicle turns anywhere; moreover, complex maneuvering involving intersections is only referred to in the context of higher automation levels. We interpret this removal to indicate that J3016:2016 narrows the scope of lateral movement to intentionally exclude handling intersections as would be required for urban driving, which is clearly intended to be in scope for higher levels of automation. Later versions of J3016 similarly do not re-introduce turning into the scope of the Level 2-relevant DDT definition. *See* J3016:2018, *supra* note 27; J3016:2021, *supra* note 27.

and Level 1 vehicles with no special driver training (beyond cursory vehicle feature familiarization),[249]

3. Any safety-relevant vehicle behavior issue should consist of an automation error of omission by the vehicle automation system rather than an error of commission,[250] and

4. Any safety-relevant vehicle behavior should be readily mitigated by driver intervention via conventional vehicle controls, namely: exerting force on the steering wheel or depressing the brake pedal.[251]

The presumption would be that an ordinary driver with no special driving training would be able to handle a Level 2 system safely. Cursory familiarization with vehicle controls would be expected so the driver understands which features are available and how to activate them. But for a Level 2 system, specialized driving skills should not be required, nor should training in how to manage faults beyond resuming normal driving or otherwise overriding features naturally.[252]

---

249. Consider, for example, the orientation that would be provided to a rental car user picking up a car at a busy airport rental lot who is running late. Credit for safety cannot be taken for any optional training that in practice is likely to be skipped in such situations, or for other reasonably foreseeable misuse regarding training requirements.

250. Errors of omission are a failure to perform some action when action should be taken, whereas errors of commission are taking an incorrect action. An example error of omission is not detecting an in-lane obstacle, whereas an example error of commission would be suddenly veering across the road centerline. This is consistent with the J3016 characterization of Level 2 having an *incomplete* OEDR rather than a defective OEDR capability. *See* J3016:2021, *supra* note 27, § 5.3 n., at 31. Some errors of omission might still be deemed unacceptable as a practical matter because the make the vehicle prone to crashes in excess of Level 0 and Level 1 mishap rates.

251. A "big red button" shutdown switch might be provided as an extra measure of safety, but safety must not rest on drivers being able to perform control actions that are not already natural responses in a conventional vehicle. This is no different than the safety approach taken to cruise control override capabilities, except also involving steering wheel takeover.

252. Examples of a natural process for overriding features would be an extrapolation of ordinary cruise control interfaces. If the human driver activates the brakes or exerts steering control, the automation features should get out of the way and let the driver take control while unambiguously annunciating to the driver that control has been ceded and the automation feature has been deactivated.

A Level 2 driver would be expected to understand performance limitations as part of the familiarization. Examples of performance limitations that might be deemed reasonable could include: gradually drifting out of lane if lane marking paint is highly degraded, failing to detect a problematic obstacle such as a haze-gray colored truck on a hill crest against a cloudy sky, attempting to drive into flood waters, or other situations in which the driver would reasonably expect the system to struggle in accordance with reasonable lay-person understanding of technical limitations of the feature.[253] Any such limitation should be communicated clearly and unambiguously to the driver as part of vehicle orientation. Automation limitations should also be consistent so that the driver can readily grasp the practical considerations of the limitations and avoid building false trust.[254]

A significant safety concern for a Level 2 system is driver "dropout"; that is, a driver paying insufficient attention to the road while not having any continuous control role in the DDT.[255] Dropout can range from unintentionally falling asleep, to becoming distracted by personal electronics devices, to daydreaming, and more. These are normal and

253. In practice applying this standard seems likely to require human subject experiments. Rather than being a disadvantage, this approach simply highlights the fundamental challenge of any Level 2 system—drivers must be good at knowing when to intervene based on their internal, subjective mental model of expectations of automation behavior. If ordinary drivers cannot build and administer a viable mental model of that type, the vehicle is unlikely to be safe in practice.

254. Arguably a vehicle that detects and avoids impact with 999 out of 1000 police vehicles when passing a police traffic stop scene is more dangerous than a vehicle that never detects such a situation. The driver of the 99.9% accurate vehicle will likely build a false expectation of perfection and not know to intervene when the thousandth police vehicle that protrudes into a travel lane is missed, potentially resulting in a crash. On the other hand, the driver of a vehicle that never detects a traffic stop will know that intervention is required every time because that situation is outside of the vehicle ODD. Whether either type of system can be designed in a way that is acceptably safe in practical use is a different question that involves DMS effectiveness as well as other issues and is beyond the scope of this article.

255. The driver drop-out phenomenon, otherwise referred to as "automation complacency" is a recurring theme in Tesla crashes and contributed to the Uber ATG testing fatality. *See* Letter from Robert L. Sumwalt, III, Chairman, Nat'l Transp. Safety Bd., Responding to NHTSA DEC. 3, 2020 ANPRM (Feb. 1, 2021) [hereinafter NTSB Response], https://downloads.regulations.gov/NHTSA-2020-0106-0617/attachment_1.pdf.

expected reactions by humans asked to perform an extremely boring supervision task.[256] Exhorting the driver to pay attention is an ineffective mitigation strategy for driver dropout. Rather, a driver monitoring system (DMS) is required to ensure that the driver pays sufficient attention to operate the vehicle safely.[257]

An effective DMS would need to have adequate safeguards to prevent reasonably foreseeable misuse, such as: exiting the driver seat while the vehicle is in motion; use of readily obtainable DMS defeat devices; a driver intentionally sleeping during a long commuting trip; and intentional operation outside the ODD, such as attempting to activate a highway-only feature on local roads or city streets.[258]

A safer approach to deploying a Level 2 system than just deploying the bare minimum required by SAE J3016[259] would be to require an effective driver monitoring system (DMS), and require enforcement of aspects of the ODD which, when violated, result in reasonably predictable misuse. Both DMS use and ODD enforcement are optional in J3016.[260] However, both are

256. It is not our intention do condone such misuse, but rather simply to point out that misuse or other failure to adequately supervise Level 2 features is inevitable due to human nature and must be accounted for in system design to achieve a level of net safety better than that for a Level 0 or Level 1 feature.

257. *See* NTSB Response, *supra* note 255, at 6–7.

258. These types of misuse have been documented and tend to be featured in online videos. An especially egregious example is a Tesla owner who was arrested for sitting in the back seat and publicly pledging to continue the misuse. *See* Simon Alvarez, *Tesla Owner Arrested Due to Autopilot Abuse Pledges to Continue Autopilot Abuse*, TESLARATI (May 12, 2021), https://www.teslarati.com/tesla-autopilot-abuser-continues-illegal-ap-use/.

259. Many, including the authors, believe that a bare minimum J3016 Level 2 system is inherently unsafe for at-scale deployment on public roads. Recall that J3016 does not purport to be a safety standard and makes no claims as to whether the defined levels might be safe. Whether a "higher functioning" Level 2 system might be deployment ready is an open question—but this very situation illustrates why J3016 is unsuitable for setting regulatory requirements.

260. *See, e.g.,* NTSB Response, *supra* note 255, at 7 (noting that NTSB recommended that NHTSA work with SAE to develop these standards).

useful to ensure acceptable safety, with an effective DMS, in particular, being widely agreed upon as a firm requirement.[261]

Any regulation that invokes SAE Level 2 should additionally require both a robust, effective DMS, and sufficient automatic ODD enforcement to deter reasonably foreseeable misuse that results in an unreasonable risk to safety.

What a driver of an acceptably safe Level 2 system would not expect to have to do is compensate for a design defect with perfect accuracy. Concrete definition of what might be a design defect in a system with incomplete OEDR capabilities can be slippery, and subject to manipulation by AV makers who might exploit unclear boundaries within J3016 to shift blame to drivers to use them as a moral crumple zone.[262]

However, some types of behaviors seem clearly unreasonable for even a Level 2 system and should be considered defects in a production system. These include:[263]

- **Sudden vehicle movements** that present a substantially increased risk of a crash. Examples of particularly hazardous movements would be a sudden attempt to cross the centerline in two-way traffic, a sudden turn toward a precipice, a sudden lateral movement toward an adjacent same-direction vehicle, and unprovoked panic stops that present risk of rear-end collisions by trailing heavy vehicles. There is a finite reaction time for any human monitoring automation, no matter how alert, and no human is perfect at reacting to unexpected vehicle misbehavior.[264]

---

261. *See, e.g.,* NTSB Response, *supra* note 255.

262. *See* Elish, *supra* note 127 (describing how responsibility for an action may be misattributed to a human actor who had limited control over the behavior of an automated or autonomous system).

263. For a compilation of risky FSD behavior (regardless of what SAE level applies), see TR (@Tweet_Removed), Thread, TWITTER (Sept. 16, 2021 and following), https://twitter.com/Tweet_Removed/status/1460999178939678724.

264. Reaction times for the full demographic range of licensed drivers must be accommodated in vehicles sold to the general public. Reaction times lengthen with age, stress, and reduced attention. Additionally, reaction time to recognize an automation failure and then respond will be longer than for reacting to a vehicle equipment failure when a human driver is in constant control. *See* NTSB response, *supra* note 255, at 6. Driver responsibilities should account for well understood limitations of human

If the vehicle misbehaves in a way that results in a mishap before a human can reasonably react or takes an action that a human cannot reasonably be expected to mitigate using normal driving controls, that is a system design flaw, not driver error.

- **DMS systems that are ineffective** at driver monitoring should be considered defective. Ineffective DMS can be expected to yield driver mental states that contribute to delayed, incorrect, or missing responses to encountering OEDR gaps in the automation system. Humans have imperfect responses to emergencies, with the probability of an unsafe response rising dramatically for highly threatening situations that require very short reaction times.[265] Every time the system misbehaves is another chance for the human driver to react incorrectly, and it is unfair for humans to be blamed for being imperfect.[266]
- **Incorrect OEDR performance** (errors of commission) violate the definition of SAE Level 2, which states that the driver completed the OEDR rather than being responsible for compensating for other-than-omissive OEDR failures: "A level 2 driving automation feature is capable of only limited OEDR, meaning that there are some events that the driving automation system is not capable of recognizing or responding to. Therefore, the driver supervises the driving automation system performance by completing the OEDR subtask of the DDT."[267]
- **Erratic OEDR performance** that causes a mishap due to violating an ordinary driver's mental model of what the automation should and should not be capable of in a way that will reasonably

performance. *See, e.g.*, Mark A. Staal, *Stress, Cognition, and Human Performance: A Literature Review and Conceptual Framework,* NASA/TM 2004-212824, (Aug. 2004), https://human-factors.arc.nasa.gov/flightcognition/Publications/IH_054_Staal.pdf.

265. *See* Staal, *supra* note 264.

266. It is important to consider typical failure rates for human intervention across the entire licensed driver demographic, accounting for both automation complacency effects not mitigated by DMS as well as poor human reaction response to sudden, high-consequence events for which drivers do not undergo continual refresher training as is the case with commercial aviation pilots.

267. J3016:2016, *supra* note 27, § 5.3 Note, at 23 (emphasis omitted).

> contribute to the occurrence of a loss event. Humans easily fall into automation complacency, and, in general, tend to trust systems more quickly than they should.[268] An automation system that works almost all the time will be trusted by the human to in practice work all the time, reducing and delaying the capability of the human driver to realize that an OEDR failure is occurring and react to it.

Overall, if a Level 2 system is net less safe than a comparably equipped Level 1 system driven in a substantively similar environment with all other aspects being equal, the presumption should be that this issue is due to Level 2 defects, and not due to excessive frequency of driver error. Any other interpretation forces the human driver into the role of a moral crumple zone.[269]

From a regulatory point of view, there should be a requirement for all SAE Level 2 production vehicles to have an effective DMS as well as ODD enforcement to mitigate foreseeable misuse to acceptable levels. Additionally, reactive regulatory measures such as recalls should be initiated for patterns of insufficient performance of DMS and ODD enforcement related to elevated occurrence rates of specific types of crashes.[270]

3. Test platform regulation involving safety drivers

Having defined the scope of a production Level 2 vehicle in terms of safety and the role of the driver, we turn our attention to the implications for testing platforms.

A testing platform for Level 2 and above automation is one that provides for automated control of speed plus steering, requires a human driver for

---

268. Patricia L. Hardre, *When, How, and Why Do We Trust Technology Too Much?*, in EMOTIONS, TECHNOLOGY, AND BEHAVIORS (2016), at 85 (noting that "[t]echnology overtrust is an error of staggering proportion, the direct and residual effects of which have become apparent locally, nationally, and internationally").

269. *See* Elish, *supra* note 127 (describing how responsibility for an action may be misattributed to a human actor who had limited control over the behavior of an automated or autonomous system).

270. The NHTSA investigation into the relationship between the use of Tesla Auto Pilot and crashes into emergency vehicles provides one example. *See* NTSB Response, *supra* note 255, at 7–8.

safety, and does not meet the requirements for a production Level 2 vehicle defined in the preceding section. An ordinary licensed driver with no specialized driving skills cannot reasonably be expected to operate such a vehicle with acceptable safety.[271]

Hallmarks of a testing platform include, but are not limited to:

- Potential for software defects that cause sudden, dangerous vehicle motion such as crossing the centerline into opposing traffic, crashing into vehicles in adjacent lanes, or turning into opposing traffic lanes. While a safety driver might be able to mitigate such risks much of the time, the need to intervene for mitigation makes a feature that displays such behaviors a test platform if trained safety drivers are being used (or, alternately, a defective deployed feature for ordinary drivers).
- Unavailability of DMS capabilities robust enough to ensure acceptable driver attention to complete the OEDR and otherwise mitigate risks not handled by automation.
- Unavailability of automatic ODD enforcement sufficient to deter reasonably foreseeable misuse without use of supplemental procedures and operational oversight. In a test environment, such misuse includes attempting to operate in ways that activate known software defects[272] even if the operation would otherwise be allowable for the ODD in the absence of the defect.
- A need to provide test platform operators with specialized training and enforce operational protocols (such as limited length driving shifts) so that they can remain attentive and react both

271. Acceptable safety is used as a positive term for absence of unreasonable risk. One example criterion for acceptable safety might be having mishap rates at least as low as for production vehicles in comparable driving circumstances and with comparable non-AV safety features.

272. An example of activating a known software defect would be driving on two-way traffic roads after a report of a test vehicle attempting to cross the centerline in a way that presented risk of collision with potential oncoming traffic. A responsible testing approach would be to either ground the test fleet or avoid such roads until a fix for the dangerous behavior can be developed, validated, and deployed. Continuing to test in a way that risks triggering known dangerous defects in a public road setting is a sign of a defective organizational safety culture. For test platforms managing such risks is done in the framework of a safety management system (SMS) specific to the testing operation.

> quickly and correctly to potentially dangerous automation behaviors.[273]

By construction, this definition specifically excludes some types of vehicles that might part of an AV development effort that nonetheless are as well behaved as production vehicles. Examples include manually driven data gathering platforms and "shadow mode" test platforms that do not permit non-production automation features to control vehicle motion.

In short, a test platform not only requires a human driver to oversee operation, but also holds that human driver accountable for compensating for one or more potential platform safety defects, self-ensuring acceptable ability to intervene, and avoiding indulgence in misuse. This is likely to be consistently achieved only via the use of specially trained test drivers and protocols.

In practice, safe public road operation of test platforms requires the use of specialized operator selection, operator training, and operational protocols. This is the scope of SAE J3018.

The SAE J3018 standard[274] provides guidance for fallback test drivers (colloquially called "safety drivers") for highly automated vehicle test operators. The scope includes classroom instruction, training, workload management, test planning, operational safety, driver monitoring, and incident response management. The general goal of conforming to J3018 is to ensure that public road testing is performed in a safe and responsible manner. A primary source of content in J3018:2020 is the Automated Vehicle Safety Consortium (AVSC) best practice on public road-testing safety.[275]

Additionally, a safety management system (SMS) is required to ensure acceptably safe operations. An SMS is a "formal, top-down, organization-

---

273. The need for such training is not problematic for a test vehicle, but rather is essential. The point here is that if drivers need special training that is a strong indication that the vehicle is a test platform.

274. J3018, *supra* note 28.

275. AUTOMATED VEHICLE SAFETY CONSORTIUM [AVSC], AVSC BEST PRACTICE FOR IN-VEHICLE FALLBACK TEST DRIVER SELECTION, TRAINING, AND OVERSIGHT PROCEDURES FOR AUTOMATED VEHICLES UNDER TEST, AVSC00001201911 (Nov. 8, 2019), https://www.sae.org/standards/content/avsc00001201911/.

wide approach to managing safety risk and assuring the effectiveness of safety risk controls. It includes systematic procedures, practices, and policies for the management of safety risk."[276] The use of an SMS ensures that testing is done in a responsible manner. Further, it allows the quick and effective use of field feedback information to mitigate risks caused by unexpected adverse events that occur during testing. A set of guidelines published by the AVSC industry group is a reasonable starting point for a public road-testing SMS, covering safety policy and objectives (SPO), safety risk management (SRM), safety assurance (SA), and safety promotion (SP).[277]

Regulators should regulate test platforms by requiring conformance to SAE J3018 and the *AVSC Testing Best Practice* (pending potential future evolution of that document to an SAE consensus standard document). Additionally, regulators should consider guidance from the American Association of Motor Vehicle Administrators (AAMVA) regarding licensing and other administrative matters related to testing.[278]

D. An Alternative to SAE J3016

Backing a definition of testing versus production for high level automation is difficult in large part because J3016 is the wrong tool for the job. Regulators should rightfully be concerned primarily with the safety of their constituents, and J3016 is not (and does not purport to be) a safety standard. Moreover, use of J3016 encourages gamesmanship in the form exploiting the Level 2 Loophole (calling a Level 4 test platform Level 2) to evade regulation. Fixing J3016 to be suitable for regulatory purposes would be a complex and lengthy task. Fortunately, there is a simpler way.

---

276. *Safety Management System (SMS)*, FED. AVIATION ADMIN., (quoting FAA Order 8000.369), https://www.faa.gov/about/initiatives/sms (last visited Mar. 18, 2022).

277. AVSC, AVSC INFORMATION REPORT FOR ADAPTING A SAFETY MANAGEMENT SYSTEM (SMS) FOR AUTOMATED DRIVING SYSTEM (ADS) SAE LEVEL 4 AND 5 TESTING AND EVALUATION, AVSC0007202107 (July 2021), https://avsc.sae-itc.org/principle-7-5896VG-46559OG.html.

278. AM. ASS'N MOTOR VEHICLE ADMIN., SAFE TESTING AND DEPLOYMENT OF VEHICLES EQUIPPED WITH AUTOMATED DRIVING SYSTEMS GUIDELINES (2nd ed., Sept. 2020), https://www.aamva.org/getmedia/66190412-ce9d-4a3d-8b6e-28c1b80e3c10/Safe-Testing-and-Deployment-of-Vehicles-Equipped-with-ADS-Guidelines_Final.pdf

We suggest that, for regulatory purposes, a highly automated vehicle be defined as any vehicle in which a computer exerts steering control that is intended to execute turns at intersections.[279] Such capabilities bring dramatically increased risk of collisions and require significantly higher levels of technological sophistication than lane following (and perhaps lane changing) more typical of highway cruising.

This results in four vehicle types for regulatory purposes:

1. **Non-automated vehicles** (at SAE Levels 0 and 1), which are regulated as conventional vehicles.[280]
2. **Low automation vehicles** (a specific subset of SAE Level 2 vehicles). These are production vehicles meeting the criteria identified for Level 2 vehicles above that are not capable of making turns at intersections. This is a subset of vehicles currently designated SAE Level 2 by manufacturers but is representative of Level 2 vehicles on the market that have a stated ODD of roads without intersections. In essence, these are "super-smart" cruise control systems that can do both speed keeping and lane keeping in highway traffic safely. An ordinary driver operates the system and is responsible for completing the OEDR by detecting and responding to out-of-ODD situations. Both effective DMS and ODD enforcement are required. Non-omissive OEDR failures are considered evidence of a vehicle design defect. Elevated mishap rates compared to non-automated vehicles—whether attributed to driver error or not—are prima facie evidence of a vehicle design defect, with the vehicle design required to accommodate expected cognitive and performance limitations of ordinary drivers, supported by DMS and ODD limitation enforcement.
3. **Highly automated vehicles** (production versions of SAE Level 3, 4, 5 vehicles that are designed for acceptable safety). These are

279. Lane changes might be permitted as an extension of lane-keeping, as might pulling to the side of a road. The criteria presented here might be extended by considering any capability that must handle designated road user crossings (vehicular intersections, marked pedestrian crossings, non-signalized rail crossings, and bikeway crossings) also be considered a highly automated vehicle.

280. Vehicles that might be considered in testing that do not automate both steering and speed are still considered non-automated vehicles rather than test platforms for this categorization.

production vehicles for which the vehicle's own human driver cannot be blamed for a crash during or adjacent to automated operation.[281] Such vehicles should be required to conform to, as a minimum, ISO 26262, ISO 21448, and ANSI/UL 4600.[282]

4. **Automation test platforms**. These are vehicles for which safety is assured by the actions of a test driver beyond the scope expected for a production low automation vehicle. This includes a requirement for skills beyond those that can be reasonably expected from a civilian driver demographic and corresponds to the hallmarks identified previously for test platforms. Safe operation of automation test platforms might require special licensing, but regardless require special driver selection, driver training, operational protocols, and use of an SMS. Operation of such vehicles should be required to conform to SAE J3018 and the AVSC SMS guidelines.

Non-automated vehicles do not have vehicle automation capable of sustained control of the steering wheel. Other automation categories are distinguished by the role of the driver: ordinary driver (low automation), no driver responsible for safety during automated driving (highly automated vehicles), and a specially trained test driver conforming to special policies and procedures (automated test platforms).

CONCLUSIONS

Based on our analysis, the AV industry is failing to pursue an optimal strategy to create public trust—despite the industry correctly identifying the importance of building trust to the long-term success of AV technology.[283] Industry attempts to engender trust via education and

---

281. While beyond the scope of this article, a reasonable amount of transition time must be afforded a human driver to take control after exiting automated operation regardless of the cause of the transition. This exceeds the requirements of SAE Level 3, which requires only "several seconds," J3016:2021, *supra* note 27, § 3.12 n. 3, which is an unreasonably short takeover time limit to impose on human drivers if a safe outcome is desired.

282. The NHTSA DEC. 3, 2020 ANPRM, *supra* note 136, seems an excellent starting point for such regulations, although there are no doubt many more issues to be solved.

283. Building trust is important because the public is wary of AV technology. *See, e.g.*, Megan Brenan, *Driverless Cars Are a Tough Sell to Americans*, GALLUP (May 15, 2018), https://news.gallup.com/poll/234416/driverless-cars-tough-sell-americans.aspx. The IEEE has predicted that the largest barrier to widespread adoption of AVs may have nothing

indoctrination are hampered by untrustworthy actions and attitudes. The public will not long embrace a complicated product—which can cause serious injury and death when vehicle crashes and other mishaps inevitably occur—if the public distrusts both the product and the people who make it.[284] We believe the shortcomings in the approach the industry uses to develop trust stem from an adversarial posture towards law, regulation, and disclosure.[285]

To build a durable form of trust needed for long term AV industry success, AV industry participants should shift to cooperation with lawmakers and regulators by embracing appropriate engineering standards specifically relating to safety, such as SAE J3018, and to embrace ethical principles which promote trust, such as identified in IEEE 7000, to foster ethical design for AV technology. To create greater certainty over the scope of laws and regulations, legislatures, regulators, and the AV industry

---

to do with technology but will be acceptance by the public. *See* Doug Newcomb, *You Won't Need a Driver's License by 2040*, WIRED (Sept. 17, 2012, 1:42 PM), http://www.wired.com/autopia/2012/09/ieee-autonomous-2040/.

284. The risk for a new technology, in the absence of trust, is that the public "'exaggerate[s] the harms associated with an innovation' and demand[s] significantly more severe laws than are warranted." *See* Tracy Hresko Pearl, *Fast & Furious: The Misregulation of Driverless Cars*, 73 N.Y.U. ANN. SURV. AM. L. 19, 68 (2017) (quoting Kyle Graham, *Of Frightened Horses and Autonomous Vehicles: Tort Law and Its Assimilation of Innovations*, 52 SANTA CLARA L. REV. 1241, 1256 (2012)).

285. To be sure, one can point to apparent moments of cooperation between NHTSA and industry, but they represent an exception rather than the rule (often committing industry to take an action which industry already was inclined to take). *See, e.g.*, Press Release, NHTSA, NHTSA Announces Update to Historic AEB Commitment by 20 Automakers (Dec. 17, 2019), https://www.nhtsa.gov/press-releases/nhtsa-announces-update-historic-aeb-commitment-20-automakers. The current disagreement between Cruise and the city of San Francisco, however, provides an example of the adversarial approach. *Compare* Paresh Dave, *San Francisco Agency Opposes Cruise Robotaxi Application, Citing Safety*, REUTERS (Dec. 1, 2021, 4:03 PM)), https://www.reuters.com/business/autos-transportation/san-francisco-agency-opposes-cruise-robotaxi-application-citing-safety-2021-12-01/ (describing San Francisco's safety concerns), *with* Letter from Aichi N. Daniel, Att'y, Cruise LLC, to Pub. Utils. Comm'n State California (Dec. 6, 2021)), https://www.cpuc.ca.gov/-/media/cpuc-website/divisions/consumer-protection-and-enforcement-division/documents/tlab/av-programs/phase-i-av-deployment-program-al-status/20211206-cruise-llcs-reply-to-protest-and-comments-to-application-for-driverless-deployment-permit.pdf (disputing San Francisco's concerns rather than working cooperatively to address them).

ought to abandon use of the SAE J3016 levels in statutes and regulations.[286] One way to accomplish this cooperation would be for NHTSA and the AV industry to engage in "negotiated rulemaking," a procedure that NHTSA indicated in a 2020 report might be a productive way forward.[287] It would not be the first time that NHTSA has used this procedure, though the NHTSA 2020 Report describes negotiated rulemaking as "new."[288]

The regulatory posture toward automotive technology in general has historically been lax compared to other life critical technologies (such as air and rail travel), featuring a manufacturer self-certification strategy. There has been no substantive regulation relating to the performance of software in life critical situations (such as a requirement for functional safety), and no regulatory requirement to follow the foundational automotive software safety standard (ISO 26262)[289] that is, at this point, more than a decade old. The AV industry seeks to maintain—if not further weaken—this inherited lack of proactive safety regulation for computer-based automotive functionality.

We attribute the adversarial posture of the AV industry towards law and regulation to a longing for the legal environment that existed during the development of railroads in the United States. The development of the railroads introduced a disruptive technology into the United States which transformed economic life for the better for many, while causing harm to others. The AV industry and its many investors and supporters see AV technology as similarly transformative. AV technology also will be disruptive, whether the industry generally acknowledges its disruption or

---

286. *See supra* Part V.

287. *See* NEGOTIATED RULEMAKING, *supra* note 29; *see also* NHTSA 2020 REPORT, *supra* note 29.

288. *Compare* Vehicles Built in Two or More Stages, 69 Fed. Reg. 36038, 36040 (proposed June 28, 2004) (suggesting use of negotiated rulemaking), *with* NHTSA 2020 REPORT, *supra* note 29, Part IV.

289. *See* ISO 26262, *supra* note 145 (updating ISO, *ISO 26262-1:2011 Road vehicles — Functional safety* (2011)), described in Juergen Schloeffel & Joe Dailey, *ISO 26262 – The Second Edition: What's in It… and What Isn't*, TECH DESIGN F. (Apr. 9, 2018), https://www.techdesignforums.com/practice/technique/iso-26262-the-second-edition-whats-in-it-and-what-isnt/ (describing what to expect from the 2nd edition).

not.[290] Consider the taxicab business in New York City to glimpse the disruptive potential for the evolving rideshare industry.

Some might misunderstand our critique as reflecting a Luddite dislike for AV technology. But that is not so. Despite AV technology's potential for disruption, we too see great promise and potential benefits for both safety and more accessible mobility.[291] But we also see the AV industry as its own worst enemy when it comes to managing the balance between safety and trust.

A. Comparison to the Railroad Industry

The political and social dynamics of the present day differ dramatically from the conditions at the time of the emergence of railroad technology. Then, as now, the default posture of the law was that any action not prohibited is permitted. But the public tolerance for risk and expectations of government at that time no longer prevail today. Recall that the early railroads operated without effective brakes[292]—placing the obligation on the public to pay attention and "get out of the way" when a train approached.[293] People of the time were more accustomed to calamity and exhibited more self-reliance so the attempt to place the blame for accidents on the victims had some surface appeal.

Our judgement is that citizens of today expect some combination of the federal, state, and local governments to proactively engage in risk avoidance and minimization of hazards, and to compel the industry to perform responsible risk mitigation rather than pure profit maximization. A key element of risk mitigation in all industries that use computers in a

---

290. Aurora notes the disruptive potential as a risk factor. *See* Aurora S-1, *supra* note 1, at 2–3. *See generally* Neal K. Katyal, *Disruptive Technologies and the Law*, 102 GEO. L.J. 1685 (2014).

291. Indeed, one of the authors got his start on self-driving safety in the 1990's as part of the Federal Highways Automated Highway System (AHS) project. That multi-company, multi-university project included a demo event on a closed public highway—preceding the more famous DARPA challenges by many years.

292. George Westinghouse invented the air brake to remedy this shortcoming. *See* HOLBROOK, *infra* note 308, at 290.

293. *See* WOLMAR, *infra* note 305, at 191. The early railroads also characterized those injured or killed on the tracks as trespassers or vagrants. In fact, hundreds of children were killed annually while playing on tracks.

safety critical way, including rail[294]—but not yet automotive—is a requirement to follow their own industry safety standards.

In response to railroad technology, the legal system developed common law rules restricting the scope of tort liability.[295] Courts expressly acknowledged that the logic of the rules facilitated the development of the railroads and prevented the transfer of wealth from entrepreneurs to individuals who suffered harm.[296] Today, doctrines like "proximate cause" no longer provide the same protection from liability as in times past.[297] While still protective of business interests, the law provides more avenues for recovery by aggrieved plaintiffs. But the mere threat of increased tort liability does not appear sufficient to ensure proper attention to safety standards needed to obtain and maintain trust.

We can learn two important lessons from the history of railroad development and regulation that apply to the AV industry. Both relate to the corporate management of public relations.[298]

---

294. In one of the author's investigations on the rail industry, he has observed that rail systems universally conform to international safety standards, typically IEC 50126, IEC 51028 and IEC 50129 for signals, safety equipment, and automated train controls.

295. "Everywhere after 1870, negligence was proclaimed to be the general rule of the common law. In case law, the most powerful recognition of the triumph of the negligence principle can be seen in two leading cases decided in 1872-1873 rejecting strict liability principles laid down in the English case of *Rylands v. Fletcher* (1868)." MORTON J. HORWITZ, THE TRANSFORMATION OF AMERICAN LAW, 1870–1960: THE CRISIS OF LEGAL ORTHODOXY 13 (Oxford Univ. Press 1992).

296. "Under strict liability, enterprises, especially railroads, would be held liable for all injuries regardless of fault. Many jurists, including Holmes, devoted themselves to marginalizing this feared authority for redistribution in torts." *Id*.

297. *See, e.g.*, Ryan v. N.Y. Cent. R.R. Co., 35 N.Y. 210 (1866) (protecting a railroad from liability for a fire its equipment clearly caused by application of the "proximate cause" doctrine).

298. Many commentators identify the railroad industry as the place where companies first began to orchestrate public relations efforts in a serious and organized way. *See, e.g.*, Nneka Logan, *The Rise of the Railroad in Virginia: A Historical Analysis of the Emergence of Corporate Public Relations in the United States*, 7 PUB. RELS. INQUIRY 5 (2018); Mark Aldrich, *Public Relations and Technology: The "Standard Railroad of the World" and the Crisis in Railroad Safety, 1897-1916*, 74 PA. HIST.: J. MID-ATLANTIC STUD. 74 (2007).

First, in the case of the railroads, the public felt abused by monopoly pricing of freight terms.[299] (The public also eventually became concerned by the abysmal safety record for railway workers, with concerns over passenger safety a distant third.) The railroad robber barons let pricing get out-of-hand and, though it took time, effective national regulation eventually followed with the formation of the Interstate Commerce Commission (ICC).[300]

The regulatory pendulum then swung the other way, with the ICC approving almost no rate increases despite a desperate need to upgrade infrastructure.[301] Moreover, the ICC's jurisdiction was expanded from rate setting to safety regulations (such as a proposal to require the ratchet wheel on hand brakes to have no fewer than 14 teeth).[302] The ICC (together with reformers and the trade press) proposed use of block signals and steel passenger cars, even suggesting control over railroad operating practices.[303] Though not every proposal found its way into legislation, the result was a rail system in decline, as we see today.

The lack of public trust in the railroads due to high freight rates created fertile conditions for the creation and expansion of this harsh regulatory scheme. In a state of distrust over rate setting, when the rail industry experienced a sharp increase in accidents (due, in part, to increased usage),

---

299. Popular dissatisfaction with freight pricing by railroads originated with the Grange movement. *See* Lenny Frank, *Farmer's Rebellion: A History of the Grange Movement*, HIDDEN HISTORY (Feb. 8, 2017), https://lflank.wordpress.com/2017/02/08/farmers-rebellion-a-history-of-the-grange-movement/. *See also* WOLMAR, *infra* note 305, at 205.

300. HOLBROOK, *infra* note 308, at 295.

301. For example, in 1917 railroads appealed to the ICC for a 15-percent rate increase to help offset rising costs associated with wartime traffic and raise revenue to invest back into network enhancements—a request the ICC rejected. The railroad industry was not deregulated until the Staggers Rail Act of 1980. This act was intended to address the concern that if the rail industry could not substantially increase its rate of return it would be nationalized. *See* Clifford Winston, *The Success of the Staggers Rail Act of 1980*, AEI-BROOKINGS JOINT CENTER FOR REGULATORY STUDIES (Oct. 2005) (noting that rail's regulated rate structure contributed to the industry's decline), at 2.

302. Aldrich, *supra* note 298, at 80–82.

303. Aldrich, *supra* note 298, at 82–83 (noting ICC's request to take over railroad operating practices).

the "result was an avalanche of proposals for what the carriers viewed as intrusive and expensive safety regulations."[304]

This illustrates one downside of taking an adversarial approach to regulation (in the railroad case, taken in the blind pursuit of profits) and continuing an adversarial approach to safety measures in the face of mounting accidents. Most would agree that neither the rail industry nor the public benefitted from this regulatory scheme. The AV industry would do well to avoid conditions that might lead to a similar destructive overregulation backlash for autonomous driving systems. One proactive measure would be to support legislation which made it mandatory for an AV company to follow its own industry standards, perhaps those agreed via a negotiated rulemaking process.

Second, though it took a while for the public to become concerned with passenger safety in rail travel, the motivation for safety improvements followed certain high-profile accidents before the Civil War.[305] The intervention of the Civil War temporarily diverted public attention away from train accidents. But public focus again turned to railroad safety "fueled by a series of yet more eminently preventable accidents."[306]

---

304. *See* Aldrich, *supra* note 298, at 81. Congress passed the Accident Reports Act in 1901, followed in 1907 by mandated hours limitations for train workers and railroad telegraphers, and required locomotive ashpans that could be cleaned without the need for a person to go underneath a train (passed in the face of significant carrier opposition). *Id.* at 77, 81.

305. In early 1853, the eight-year-old son of President-elect Franklin Pierce died in a rail accident. CHRISTIAN WOLMAR, THE GREAT RAILROAD REVOLUTION: THE HISTORY OF TRAINS IN AMERICA (2012), at 193. Later at Grand Crossing in Chicago, on April 25, a train failed to observe a stop signal, causing the death of 21 immigrant German passengers. *Id*. at 194. Less than two weeks later, on May 6, an engineer ignored a stop signal at a drawbridge, and the train plunged into the Norwalk River in Connecticut resulting in 46 deaths. *Id*. These incidents were in addition to an increasing number of head-on train collisions termed "cornfield meets." *Id*. The world's worst railroad accident to date took place in 1856 at Camp Hill on the North Pennsylvania Railroad when 56 fatalities took place in the "Picnic Train" disaster which involved a Sunday excursion train. *Id*. at 195. The two-train collision also injured 100 people. "At root the problem was that the rapid spread of the railroads had not been matched by the technological changes required to keep them safe." *Id*. at 196. Accidents of this magnitude might occur in an AV application used to platoon trucks.

306. *Id*. at 196.

Nevertheless, the railroad industry obstructed deployment of safety devices, seeing accidents as "an unfortunate but unavoidable side effect of an industry in which large machines moved at high speed."[307]

Indeed, the automatic coupler and air brake gained universal adoption because of a concerted effort led by Lorenzo Coffin, whose incessant badgering of lawmakers to make the Westinghouse air brake and the Janney coupler mandatory led to the Railroad Safety Appliance Act, which became fully operative by 1900.[308] Even then, the railroads objected because, they argued, it was a mistake to mandate the use of a particular technology (a common argument across industries to avoid regulation).[309]

The public accurately perceived that the railroads had not embraced safety measures. This likely contributed to a further decline in the reputation of the railroad industry, damaging an image already tarnished by perceived rate gouging. Our interpretation is that the laudable public relations efforts which started in Pennsylvania after 1900 proved insufficient to rehabilitate reputations once lost.[310] This left the railroads too weak effectively to challenge or reverse the stringent rate regulation that contributed to the decline of that industry. The AV industry may find itself in a similar position if it fails to follow its own industry safety standards pursuant to an effective legislative mandate.

---

307. *Id*.

308. STEWART H. HOLBROOK, THE STORY OF AMERICAN RAILROADS 291, 295, 297 (1947).

309. WOLMAR, *supra* note 305, at 200.

310. The Pennsylvania Railroad led the way by featuring safety in its promotional literature. In 1906, the Pennsylvania Railroad hired Ivy Lee as a response to the major safety crisis that affected the railroads. "Lee's most important contribution was to stress the need for candor on safety matters." Aldrich, *supra* note 298, at 84. The passenger fatality rates had increased to 24 per billion miles traveled. (By comparison, American commercial air travel in the 1990's had a fatality rate of .27 fatalities per billion miles traveled. Aldrich at 78.) In "Slaughter on Railroads", the *Chicago Daily News* asserted that passengers and trainmen were being butchered day by day because railroads found it cheaper to kill than not to kill. Aldrich at 79. At Lee's urging, the Pennsylvania Railroad made the results of its crash investigations public. Aldrich at 84. The AV industry has the time to take proactive steps to avoid these types of headlines and all the harm that might follow by simply truly embracing a safety culture and a commitment to following industry standards.

B. Recommendations to the AV Industry

Currently, the AV industry is unquestionably anti-regulation. This risks further degradation of trust every time there is a high-profile adverse media event.[311] An industry strategy of painting whichever company is implicated in the fatality as a bad, rogue actor will only hold up for so long. The industry is taking a large gamble that it will be able to deploy AVs in a convincingly positive light before an increasing number of adverse events— and the consequent public outcry—forces regulators' hands. As the promised deployment timeline for AV technology drags on year after year (while the industry currently has plans for near term deployment of AVs at scale), this seems an increasingly risky strategy.

But whatever may come, we are confident that the AV industry would do well to embrace the reality of a safety culture rather than merely propound the illusion of safety in a technological Potemkin village. We fear that the latter approach, while potentially expedient in the short term, will prove unwise in the long term. It is in that spirit that we offer up our observations about the importance of trust and suggest one approach to build lasting trust by truly embracing a culture of safety—starting by embracing safety standards drafted by the AV industry itself and tailored in a cooperative negotiated rulemaking process to meet the needs of industry, regulators, and the public.

---

311. In this article, we make a case for regulation of AV technology by focusing on how the AV industry might lose trust based on actions industry participants take (or do not take) such as: failing to comply with law or published industry standards, failing to disclose deployment standards, and perpetrating myths which conceal capabilities and manipulating democratic processes. Trust in AI, however, requires more than taking the actions we suggest. Further ethical questions remain about the specific capabilities of AI systems and their use relating to privacy, the potential for automated data usage to discriminate against protected classes, and the status of an AI system (or "robot") as a moral agent. *See, e.g.*, Benjamin Kuipers, *Perspectives on Ethics of AI: Computer Science*, *in* OXFORD HANDBOOK OF ETHICS OF AI 421 (Markus Dubber, Frank Pasquale, & Sunit Das eds., Oxford U. Press, 2020) (noting how an AV as a moral agent must do more than choose the lesser evil when confronted by an ethical dilemma –such as a choice of which pedestrian to hit in an unavoidable collision–but instead recognize upstream decision points that avoid the dilemma entirely).