Trust, Control, and Communication in the Transport Network

John D. Lee

University of Wisconsin—Madison

jdlee@engr.wisc.edu

Trust in Automation

- *"Trust ...* the attitude that an agent will help achieve an individual's goals in a situation characterized by uncertainty and vulnerability." (Lee & See, 2004)

- But what about:

  Communicating through performance?
  Control of "Autonomous" systems?
  Incidental users?
  Networks of people and technology?

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46*(1), 50–80.
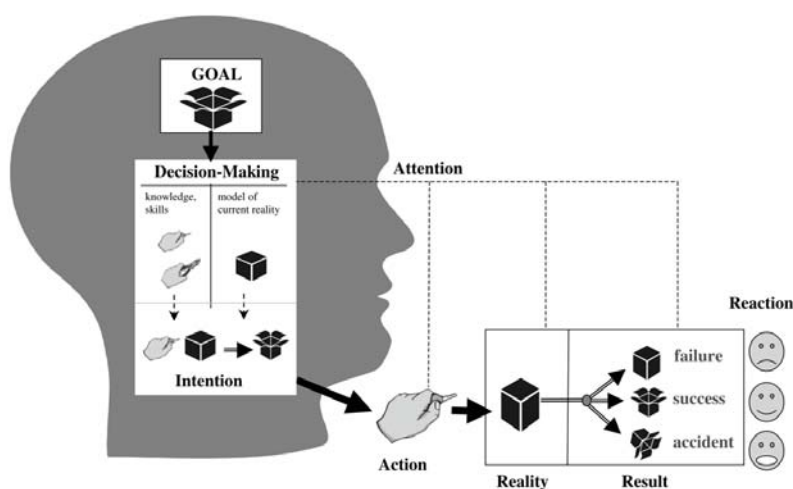

Singular and Networked Automation
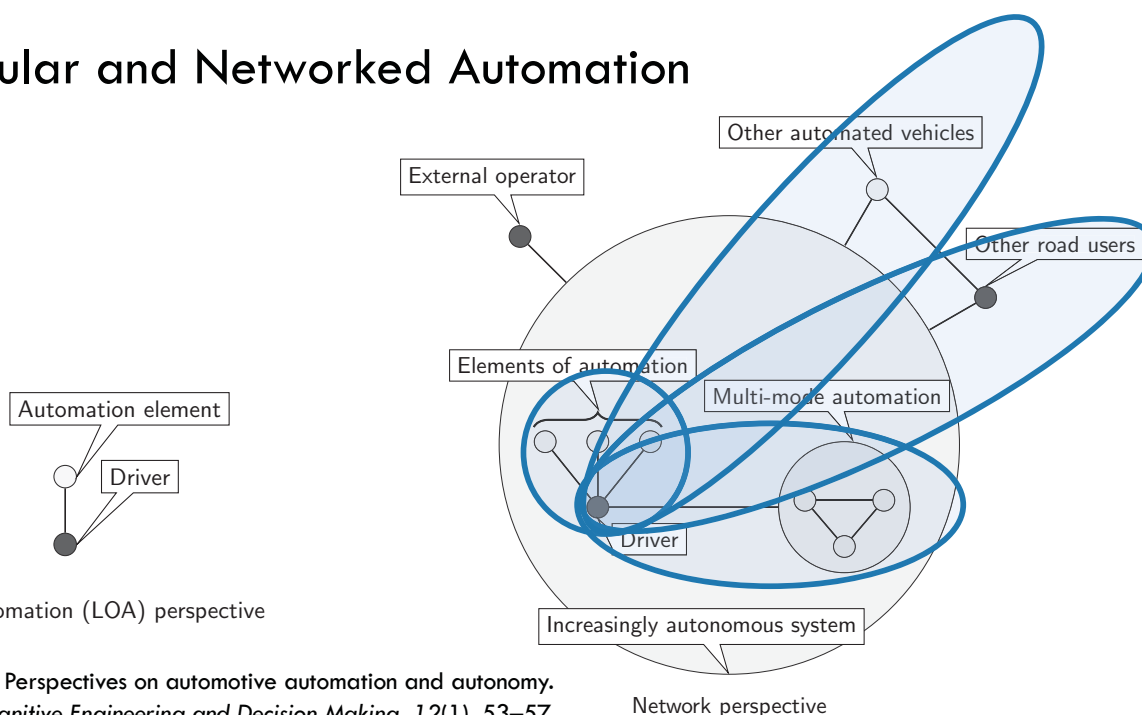
Automation element

Driver

Levels of automation (LOA) perspective

Lee, J. D. (2018). Perspectives on automotive automation and autonomy. *Journal of Cognitive Engineering and Decision Making, 12*(1), 53–57.

External operator

Other automated vehicles

Other road users

Elements of automation

Multi-mode automation

Driver

Increasingly autonomous system

Network perspective

Varieties of Control

• Pragmatic—Control to achieve goals

• Epistemic—Control to learn

• Communicative—Control to signal intent to others

• Existential—Control for self-efficacy

"These two lanes will take us anywhere" (Springsteen, 1975)
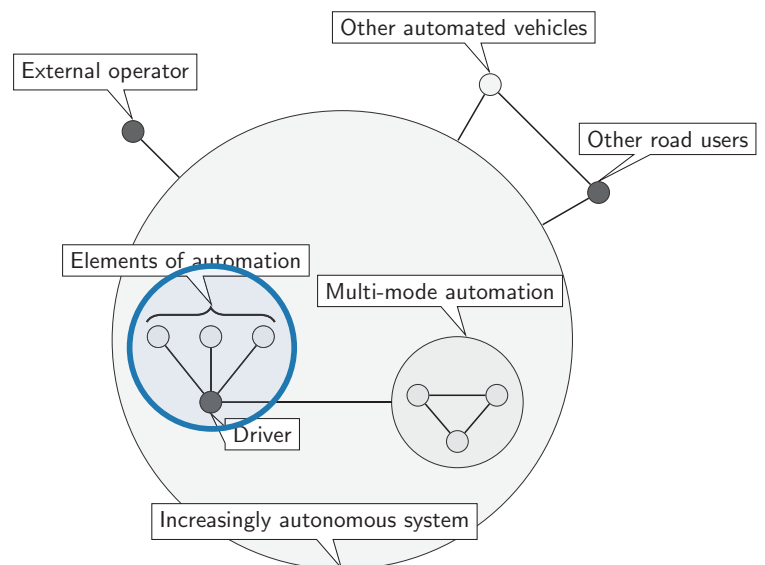

Control: From intention to action

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *The Behavioral and Brain Sciences, 28*, 675-691

Control:
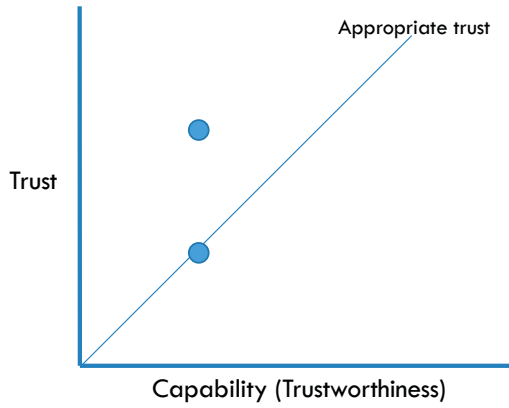# Communicating intention and coordinating joint action
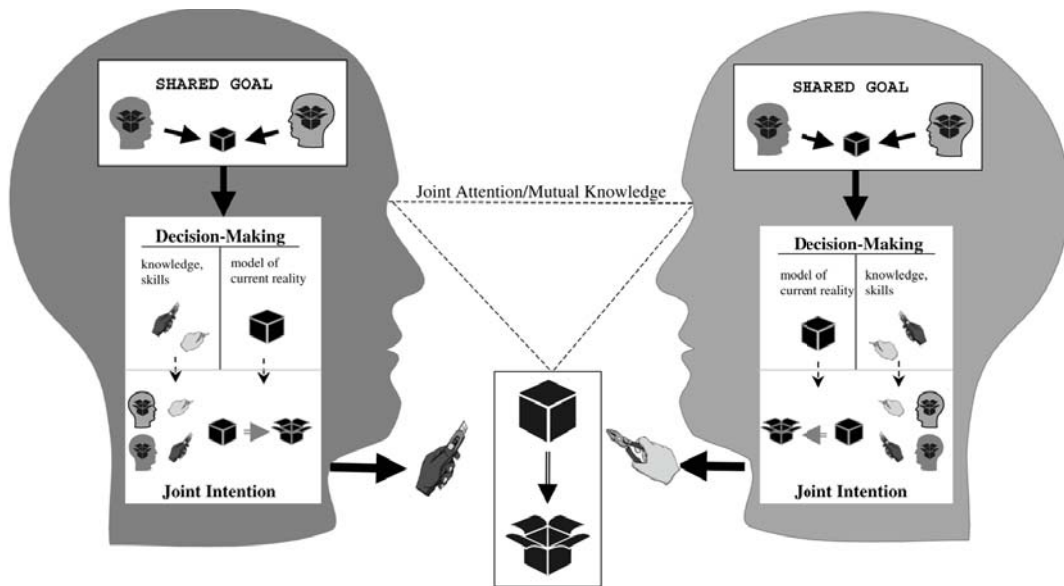
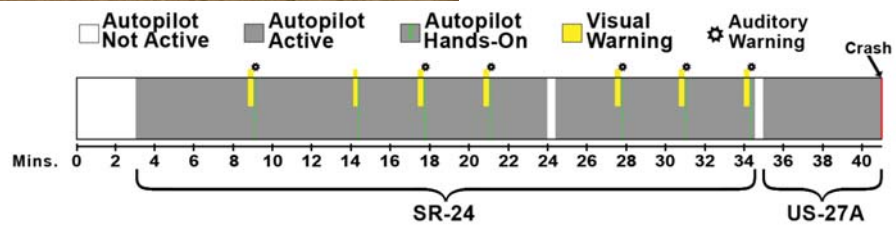# Trust in sometimes-semi-self-driving cars

Lee, J. D. (2018). Perspectives on automotive automation and autonomy. *Journal of Cognitive Engineering and Decision Making*, 12(1), 53–57.

Network perspective
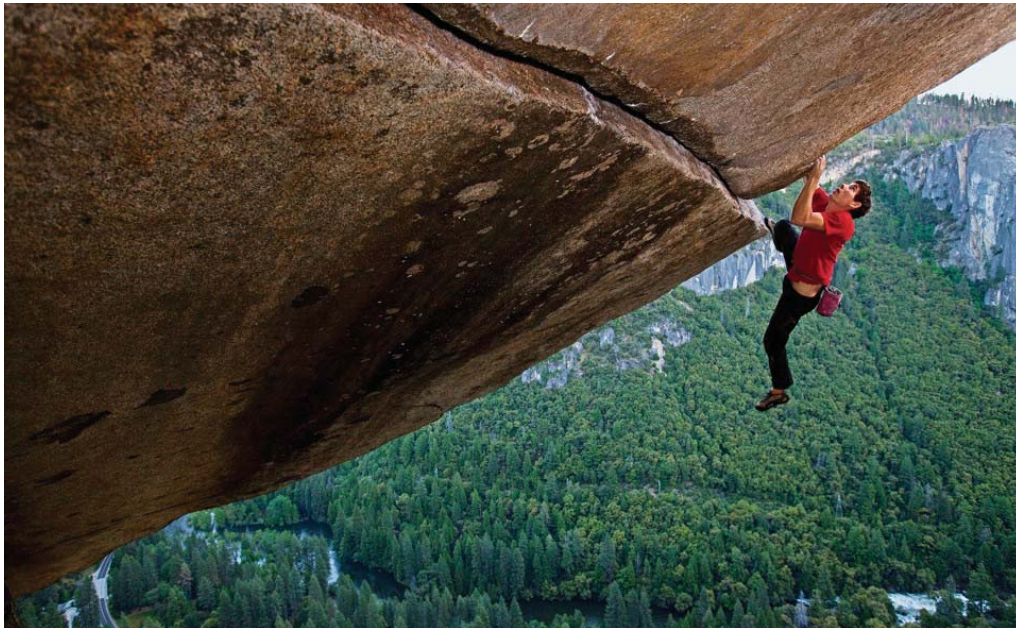
Overtrusting Automation

Trust vs. Capability (Trustworthiness) graph showing Appropriate trust line

NTSB Investigation of Fatal Tesla Crash

Autopilot Not Active | Autopilot Active | Autopilot Hands-On | Visual Warning | Auditory Warning | Crash

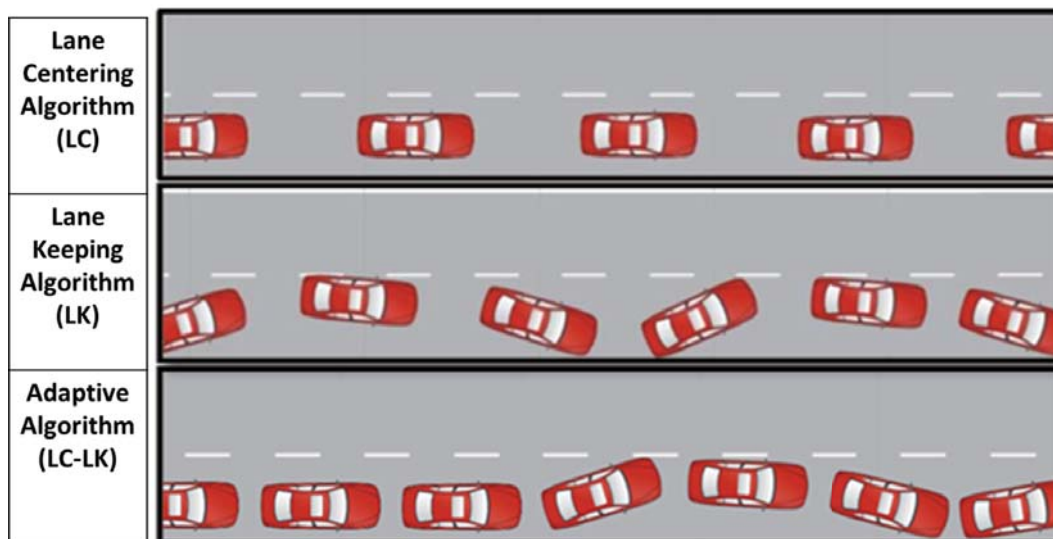Mins. 0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40

SR-24   US-27A

Risk is Not So Obvious with "Self-driving" Cars

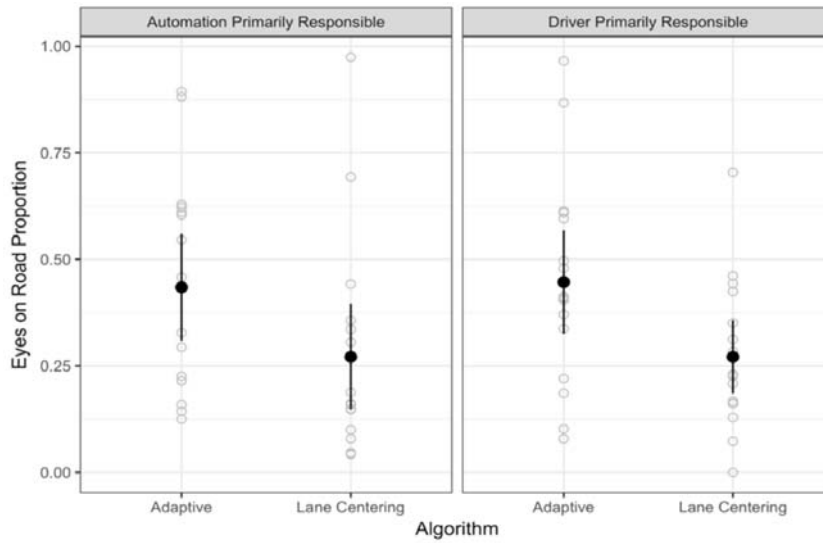

Algorithms to Mitigate Overtrust

Price, M. A., Venkatraman, V., Gibson, M. C., Lee, J. D., & Mutlu, B. (2016). Psychophysics of trust in vehicle control algorithms. *SAE Technical Paper*, (2016-01–0144).

Algorithms to Mitigate Overtrust

Alsaid, A., Lee, J. D., & Price, M. (in press). Moving into the loop: An investigation of drivers' steering behavior in highly automated vehicles. *Human Factors*.

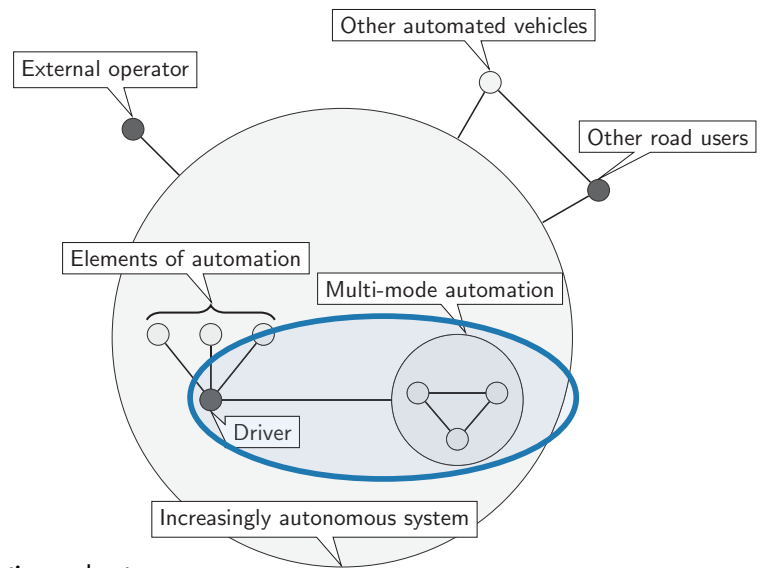# "Vicarious Steering" Indicates Trust
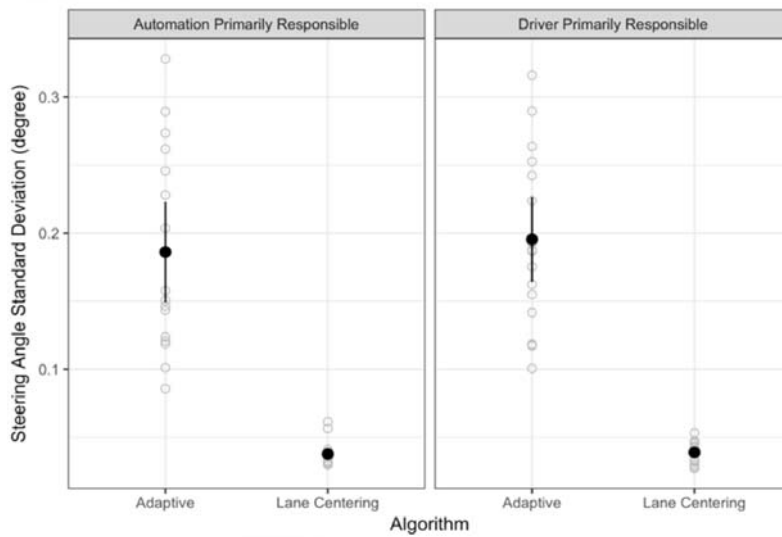
Alsaid, A., Lee, J. D., & Price, M. (in press). Moving into the loop: An investigation of drivers' steering behavior in highly automated vehicles. *Human Factors*.

Trust in Self-Driving Cars

Lee, J. D. (2018). Perspectives on automotive automation and autonomy. *Journal of Cognitive Engineering and Decision Making, 12*(1), 53–57.

Network perspective

Undertrust and Comfort in Driverless Cars
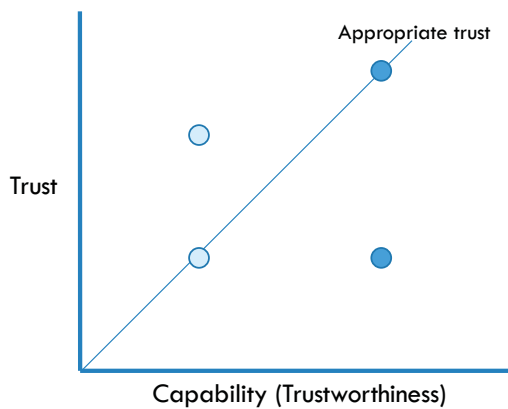
Determinants of trust:
Micro—moment-to-moment movement
Macro—Reputational and societal influence

Drivers Press the Brake When There Is No Sign
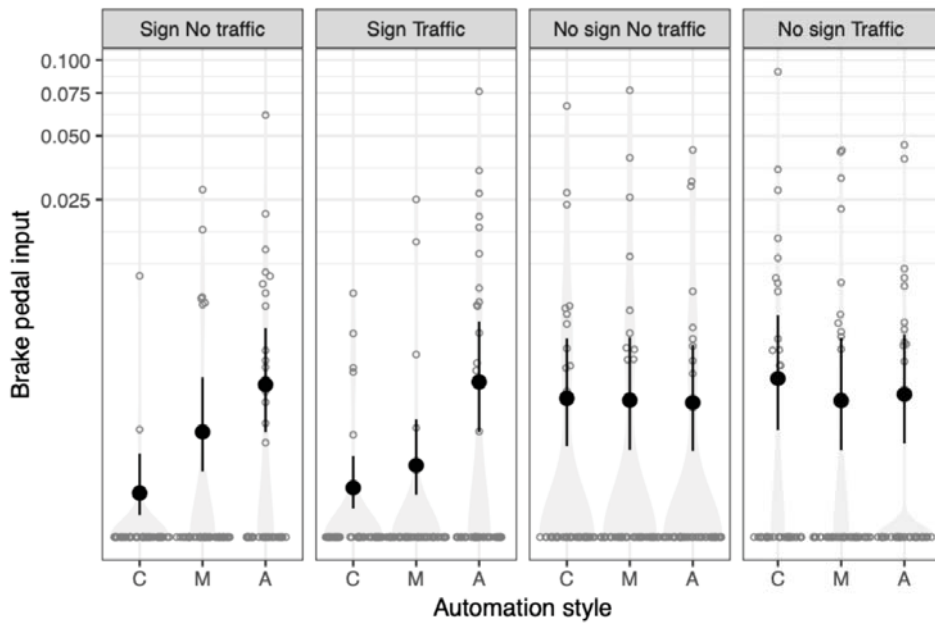
Drivers Press the Accelerator After Experiencing Slow Intersections
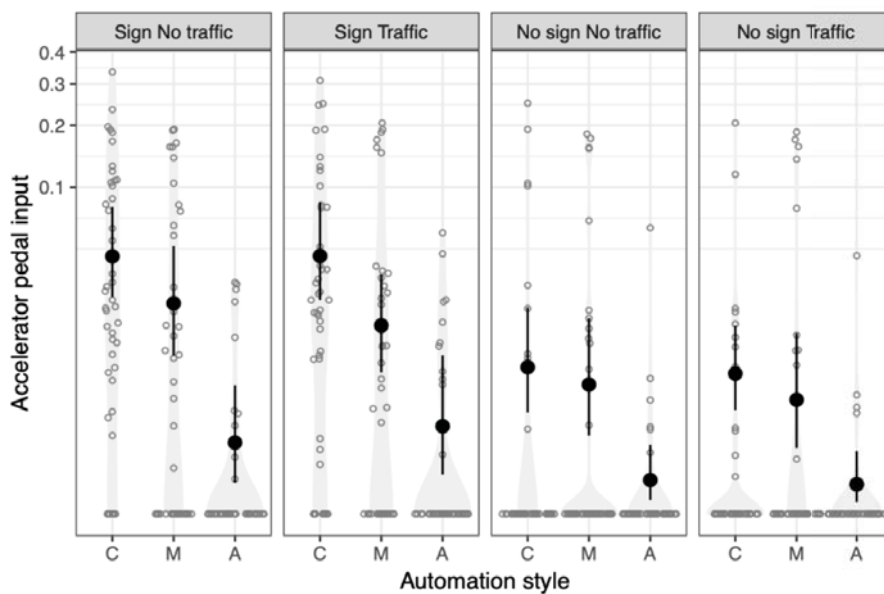
How to Model Accelerator Input?
Discrete response the continuous

Two-part Model: How often and how much
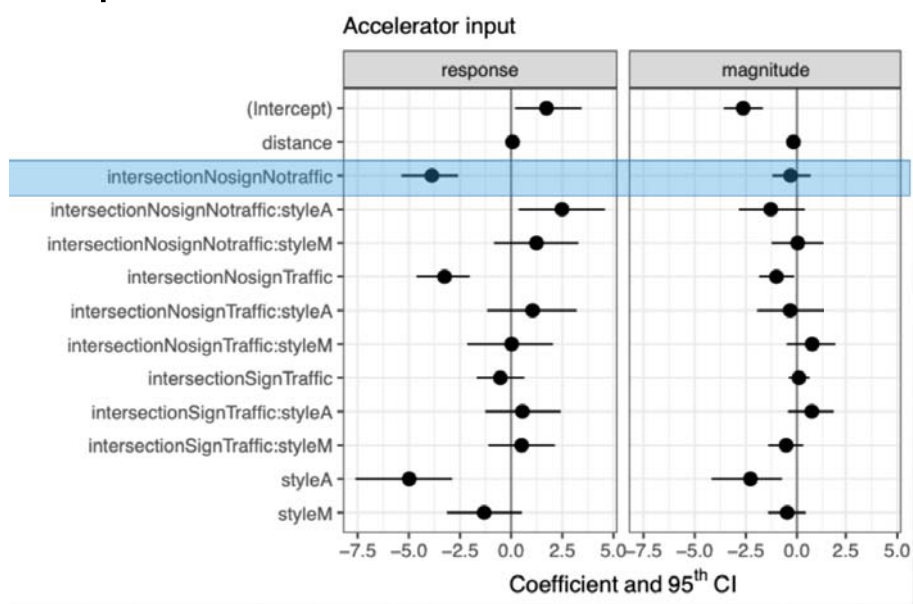
Lee, J. D., Liu, S.-Y., Domeyer, J., & Dinparastdjadid, A. (in review). Assessing driver acceptance of fully automated vehicle with a two-part model of intervention tendency and magnitude. *Human Factors*

Trust in Driverless Cars
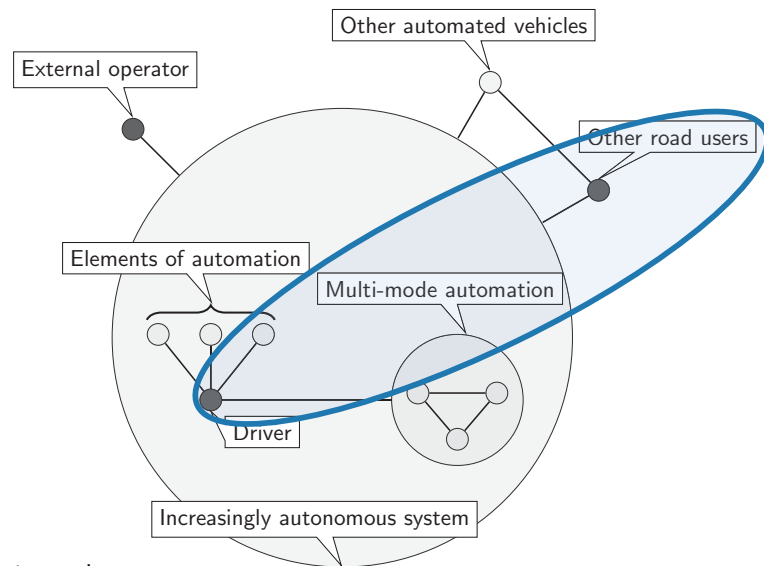
Lee, J. D. (2018). Perspectives on automotive automation and autonomy. *Journal of Cognitive Engineering and Decision Making, 12*(1), 53–57.

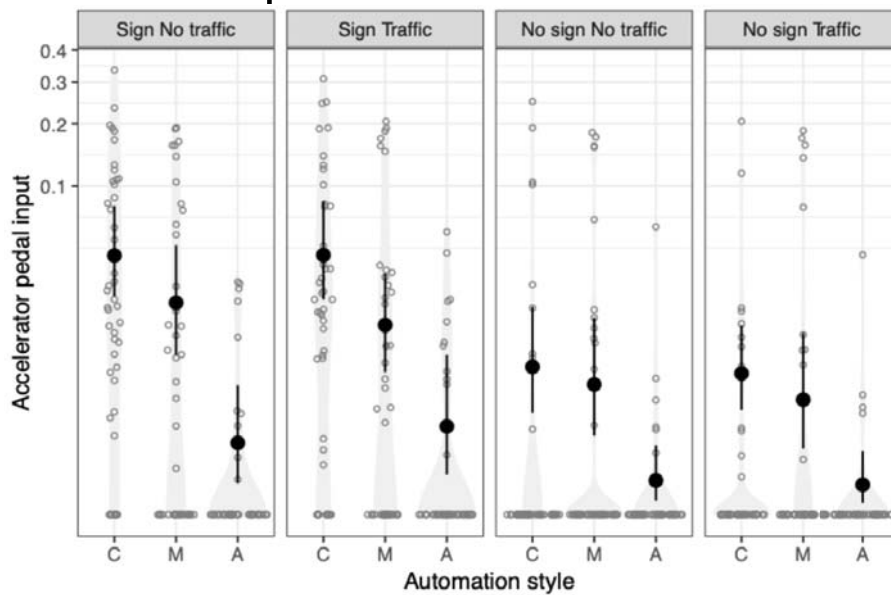# Micro Determinants and Consequences of Trust: Chronemics, Proxemics, and Kinesics

Doymeyer, J., Dinparastdjadid, A., Lee, J. D., & Douglas, G. (2019). Proxemics and kinesics in automated vehicle-pedestrian communication: Representing ethnographic observations. *Transportation Research Record: Journal of the Transportation Research Board.*

Trust in Driverless Cars

External operator

Other automated vehicles

Other road users

Elements of automation

Multi-mode automation

Driver

Increasingly autonomous system

Network perspective

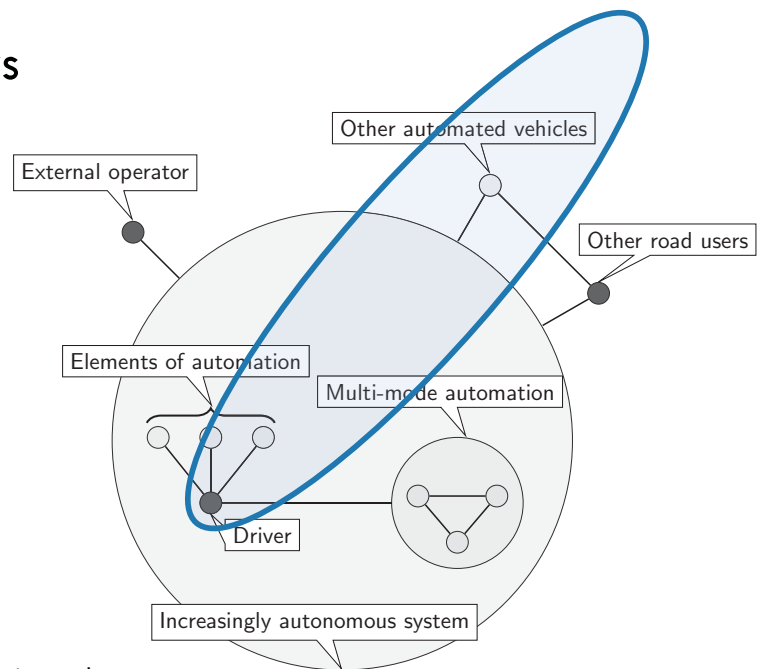Lee, J. D. (2018). Perspectives on automotive automation and autonomy. *Journal of Cognitive Engineering and Decision Making*, *12*(1), 53–57.

Macro Determinants and Consequences of Trust

J.D. POWER

J.D. Power Studies

U.S. Tech Choice Study

- Online panel survey 7,947 in 2016 and 8,517 in 2017

- "How much would you trust the ability of a vehicle equipped with self-driving technology to operate without a human driver's input?"

- Open-ended responses with Likert ratings

Lee, J. D., & Kolodge, K. (2018). Understanding attitudes towards self-driving vehicles: Quantitative analysis of qualitative data. *Proceedings of the Human and Ergonomics Society Annual Meeting*, *62*(1), 1399–1403.

Responses Are Less Favorable in 2017: Why?
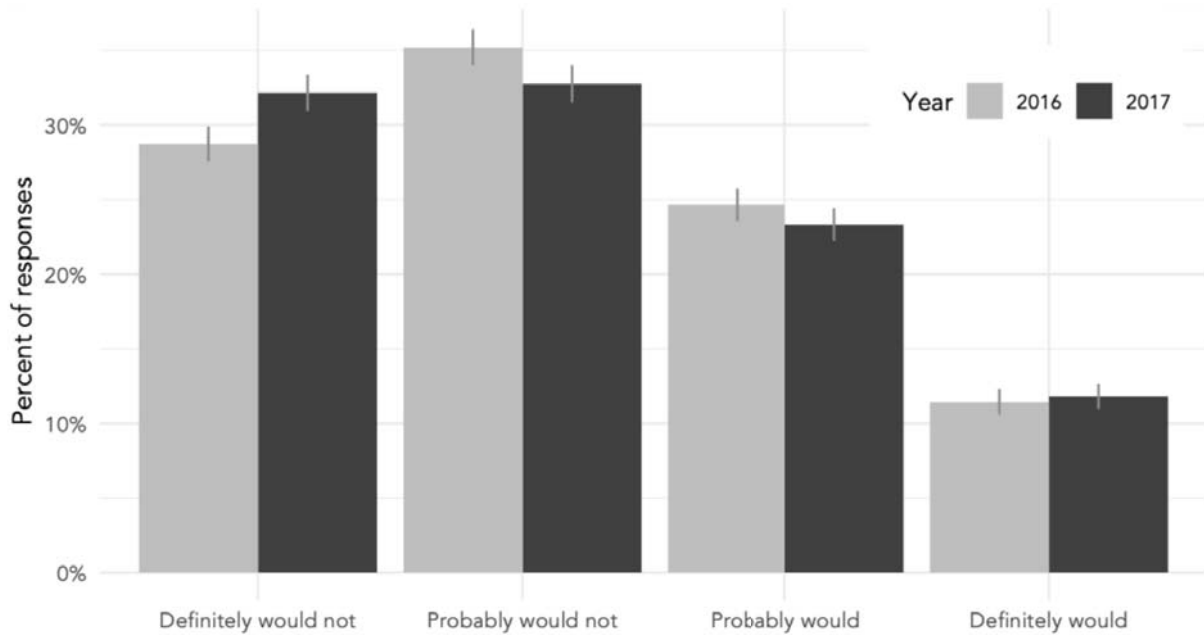
Topic Modeling for Quantitative Analysis of Qualitative Data
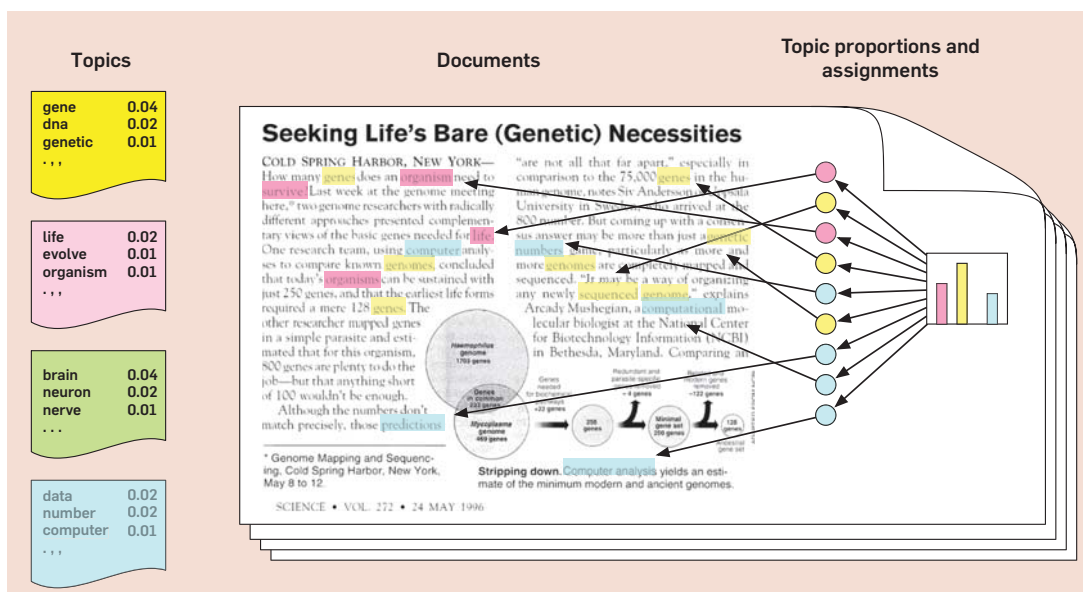
Blei, D. M. (2012). Introduction to probabilistic topic models. *Communication of the ACM, 55*(4), 77–84.

Distribution of Topics over Words and Comments

| Topic and keywords	Exemplar comments
Control until proven FREX: proven, issues, given, worried, safe, control	"I do not like giving up control and do not think the tech has progressed sufficiently at this point to be safe and secure" "not proven enough yet,, would feel out of control"
Feel uncomfortable FREX: feel, operation, functions, actions, risky	"I can't explain it.....I just don't feel comfortable with a car operating on its own. I like to be in charge of the car."

Distribution of Topics over Words and Comments

Topic and keywords	Exemplar comments
Trust when mature FREX: trust, point, years, life, early	"I do not trust this type technology at this point of development" "I trust the technology will be advanced to the point that it will be safe when it's available for purchase"
Technology improving FREX: develop, interested, technology, fully, fail	"Technology is at a peak and still improving. And all new technologies MUST secure federal permission to function, and NO sane person/agency would approve the use of an unsafe technology"

Topics and Their Connections
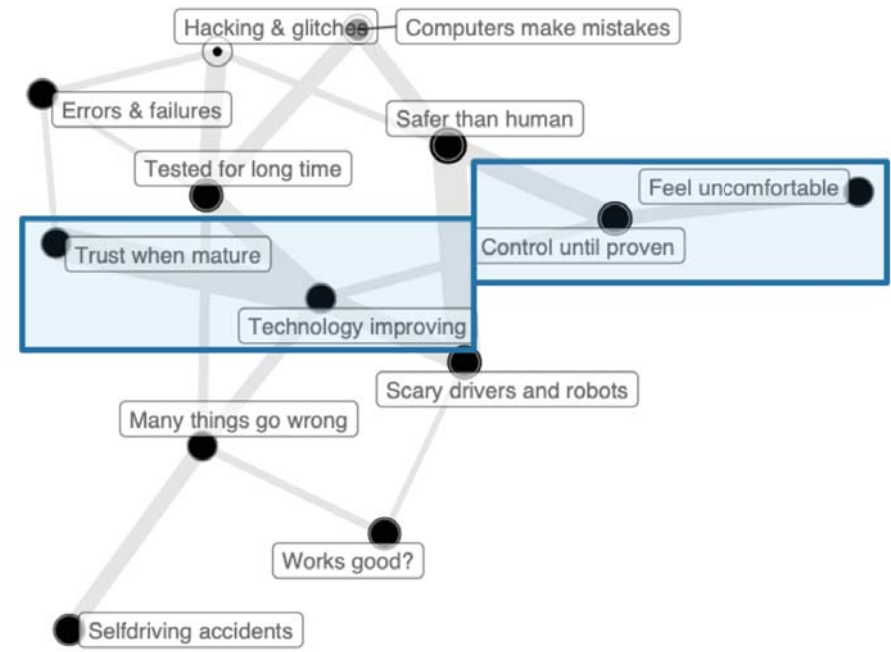
How Does Topic Prevalence Depend on Year and Rating?

- Structural topic modeling moves topic modeling from an unsupervised learning technique to a semi-supervised learning technique

- Distribution of topics across comments not assumed to be uniform, but linked to covariates (e.g., year and rating)
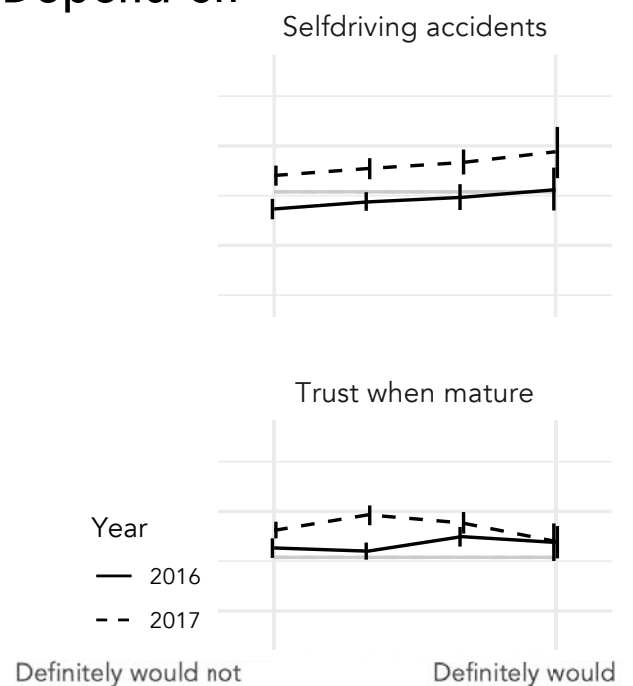
Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., … Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science*, *58*(4), 1064–1082.

How Does Topic Prevalence Depend on Year and **Rating**?

- Hacking is worrisome

- Other drivers are also worrisome

Hacking & glitches

Safer than human

Year
— 2016
- - 2017

Definitely would not   Definitely would

How Does Topic Prevalence Depend on **Year** and Rating?

- The mention of self-driving car crashes increased in 2017

- Negative comments associated trusting mature technology increased in 2017

Selfdriving accidents

Trust when mature

Year
— 2016
- - 2017

Definitely would not   Definitely would

"I can't explain it.....I just don't feel comfortable with a car operating on its own. I like to be in charge of the car." ?"

- Dread risk—Risks that are uncontrollable and unobservable

  (Slovic, P. (1987). Perception of risk. Science, 236(4799), 280–285)

> ⮌ You Retweeted
> **Khalil Sehnaoui** ✔ @sehnaoui · May 26
> When the elevator button falls off and you realize "Door Close" was always a lie.

Psychophysics of Dread Risk

- Dread risk perceived as 1000 times greater than controlled risk

  (Slovic, P. (1987). Perception of risk. *Science, 236*(4799), 280–285)

- Dread risk guides society to risky outcomes

  (Gigerenzer, G. (2004). Dread risk, September 11, and fatal traffic accidents. *Psychological Science, 15*(4))

- Trust declines with easy misses

  Madhavan, P., Weigmann, D. A., & Lacson, F. C. (2006). Automation failures on tasks easily performed by operators undermine trust in automated aids. *Human Factors, 48*(2), 241–256
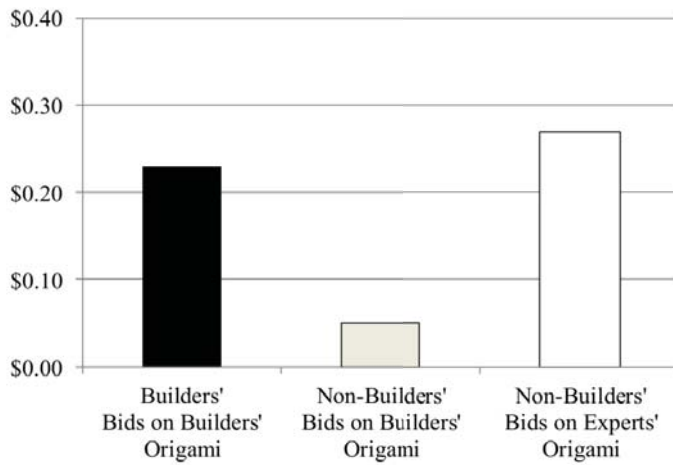

Small changes produced 100% misclassification

Evtimov, I., Eykholt, K., Fernandes, E., Kohno, T., Li, B., Prakash, A., Song, D. (2017). Robust physical-world attacks on machine learning models.

Make Less Dreadful: Make observable

Make Less Dreadful: Make directable

Lee, J. D., Wickens, C. D., Liu, Y., & Boyle, L. N. (2017).
Designing for People: An introduction to human factors engineering.

Existential Control and the Ikea Effect:
More value, less perceived risk

Nordgren, L. F., Van Der Pligt, J., & Van Harreveld, F. (2007). Unpacking perceived control in risk perception: The mediating role of anticipated regret. *Journal of Behavioral Decision Making*.
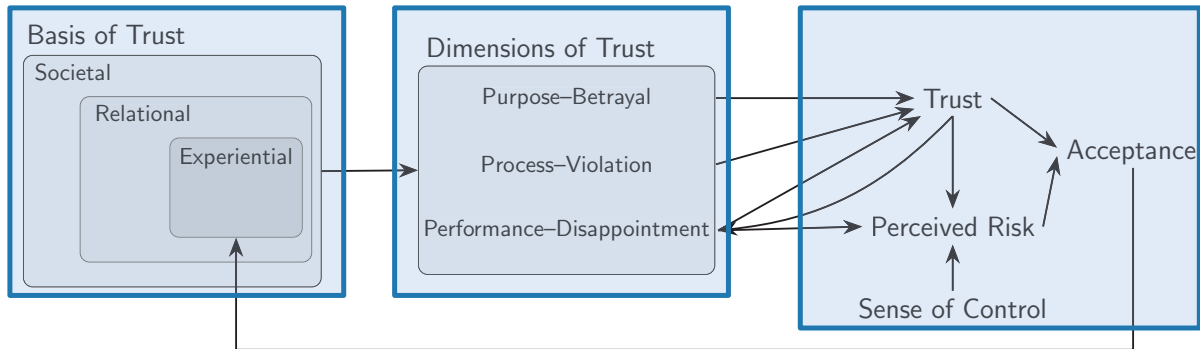
Norton, M. I., Mochon, D., & Ariely, D. (2012). The IKEA effect: When labor leads to love. *Journal of Consumer Psychology*

"I face a disconcerting void—an absence of controls"

Determinants and Consequences of Trust

- The basis of trust is not just direct experience

- Dimensions of trust shift with the basis

- Sense of control, trust, and performance combine to influence perceived risk

Lee, J. D., & Kolodge, K. (in review). Exploring trust in self-driving vehicles with text analysis. *Human Factors*.

Trust in Automation

- Communicate capability through performance

- Support control of "Autonomous" systems

- Consider trust of incidental users

- Networks influence trust
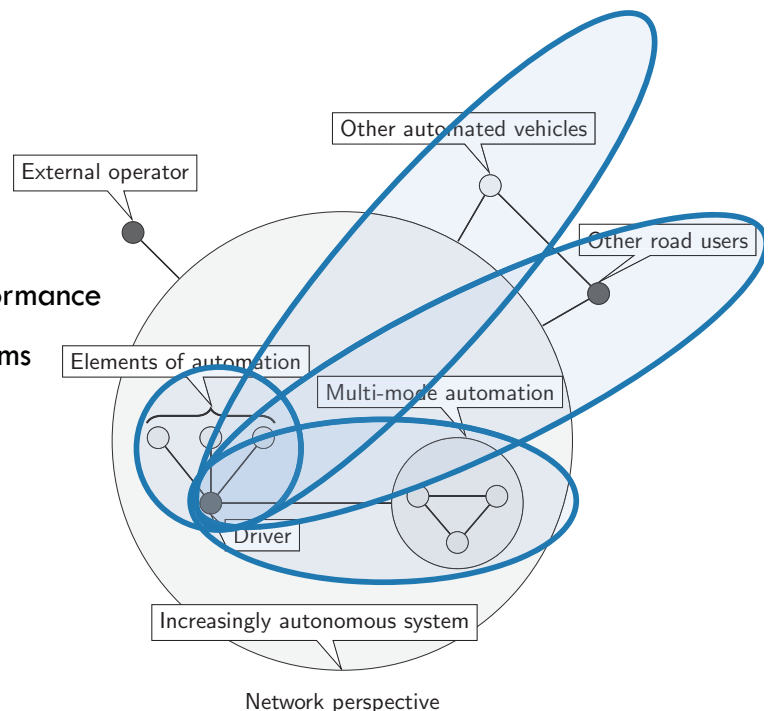
John D. Lee
University of Wisconsin
jdlee@engr.wisc.edu
Twitter: jdlee888

Network perspective