



# How Drivers Allocate Visual Attention During Extended Conversational AI Interactions

Allegra Ayala, PhD  
Senior Scientist  
Human Factors

Coauthors:

Chris Monk, PhD (Exponent)  
Christine S.P. Yu, PhD (Exponent)  
Ellie Robbins, PhD (Exponent)  
Greg Fitch, PhD (formerly Google)  
Dara Gruber, PhD (Google)

October 21, 2026

- Gemini is a family of highly capable, multimodal generative AI models developed by Google
  - Google



# Gemini in the Car

---

- Gemini lets drivers complete tasks and get answers using natural language



Android Auto

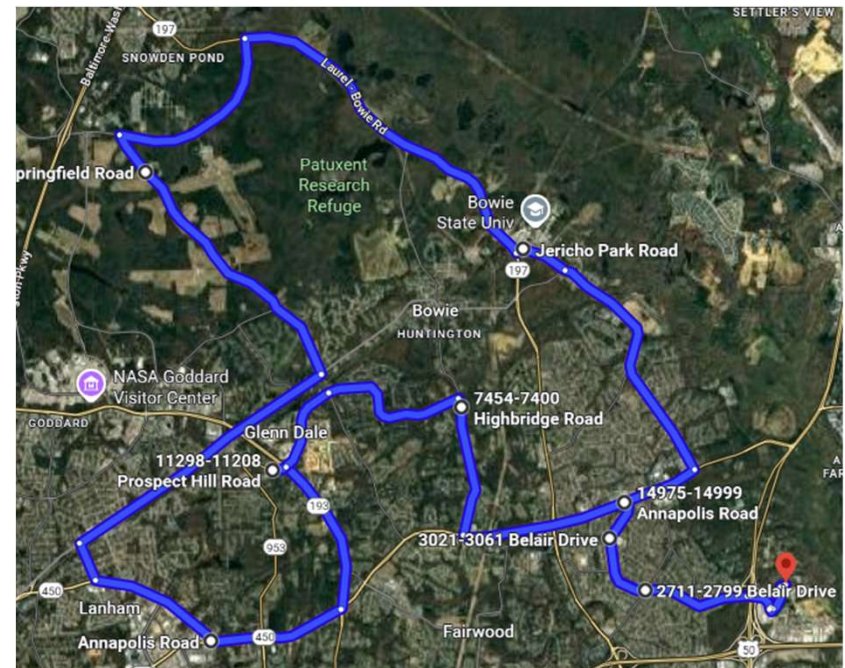
# Study Objective and Technical Approach

- Objective
  - Assess cognitive load and visual demands associated with Gemini Live in Android Auto relative to comparison tasks
- Technical Approach
  - On-road experiment measuring cognitive load and visual attention while performing Gemini Live and comparison tasks on public roads
  - Secondary analysis examined the structure and sequencing of glance episodes during conversational interactions



# Route

- Duration: 1 hour and 12 minutes
- No speeds above 45 mph (72 kmph)
  - No major highways
- Minimal stoplights
- Minimal traffic
  - Experimental start times of 9:30AM and 1:30 PM selected to minimize traffic exposure



# Tasks

---

- Driving (with Visual-manual Navigation)
- Hands-free Phone Call via Android Auto
- Gemini Live
  - One-turn
  - Multi-turn
- OSPAN



Exponent employee depicted

# Test Participants and Protocol

- 24 licensed drivers with prior voice-assistant experience participated in the study
- Pre-Drive
  - Introduced to Gemini Live with example prompts
  - Given task instructions
- Practice while Stopped
- Practice while Driving
- Testing while Driving
  - Subjective feedback after each task
- Final Subjective Feedback



Exponent employee depicted

# Measures

- Tactile Detection Response Task (DRT)
  - Miss Rate
  - Response Time
- Glance Behavior (NHTSA criteria)
  - Mean Glance Duration (MGD)
  - Proportion of Glances > 2 sec
  - Total Eyes off Road Time (TEORT)
  - **Eyes-on-road time (for Glance Episode analysis)**
- Subjective Ratings
  - Modified TLX



Exponent employee depicted

# Sample Images from Tobii Camera

On-Road Glance



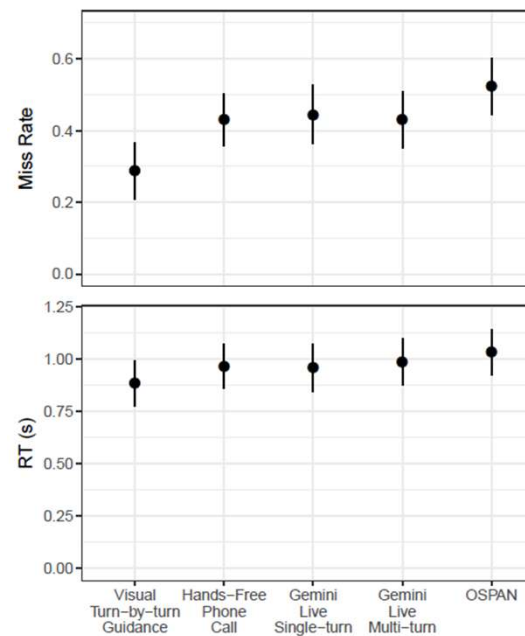
On-Device Glance



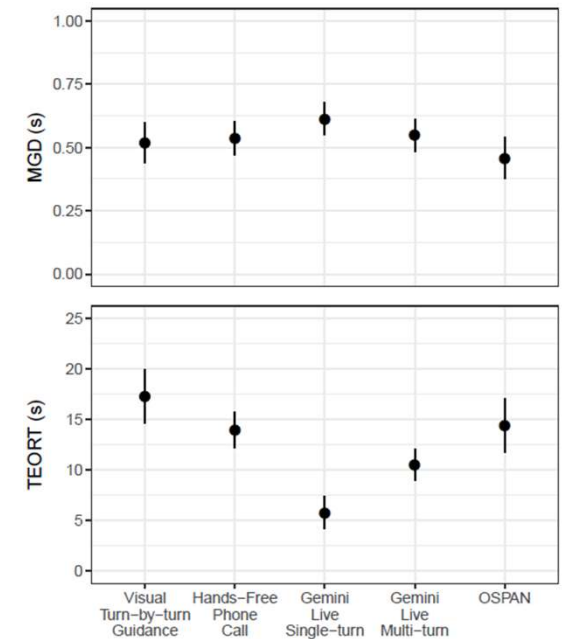
# Overall Study Findings

- Gemini Live produced cognitive workload comparable to a hands-free phone conversation and lower than OSPAN
- Visual demand remained low, meeting key NHTSA glance criteria and outperforming other tasks on TEORT
- Participants rated Gemini Live similarly to hands-free phone calls and navigation, and less demanding than OSPAN

### DRT Results



### Glance Behavior Results



# Why Analyze Glance Episodes?

---

- Traditional metrics summarize overall visual demand (e.g., mean glance duration, TEORT)
- Similar aggregate demand can arise from different patterns of visual attention allocation
- Episode-based analyses examine how drivers alternate between roadway monitoring and device-directed glances

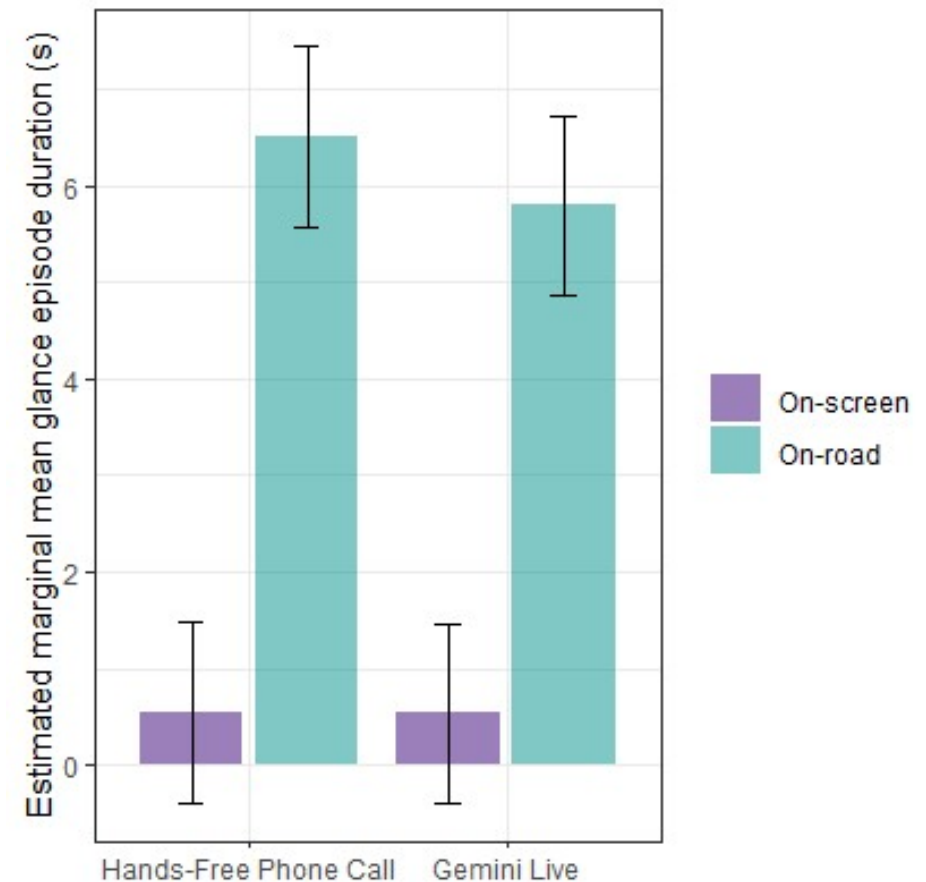
## Glance Episodes Main Finding

---

- Extended conversational AI interactions produced visual attention patterns comparable to conventional hands-free phone conversations

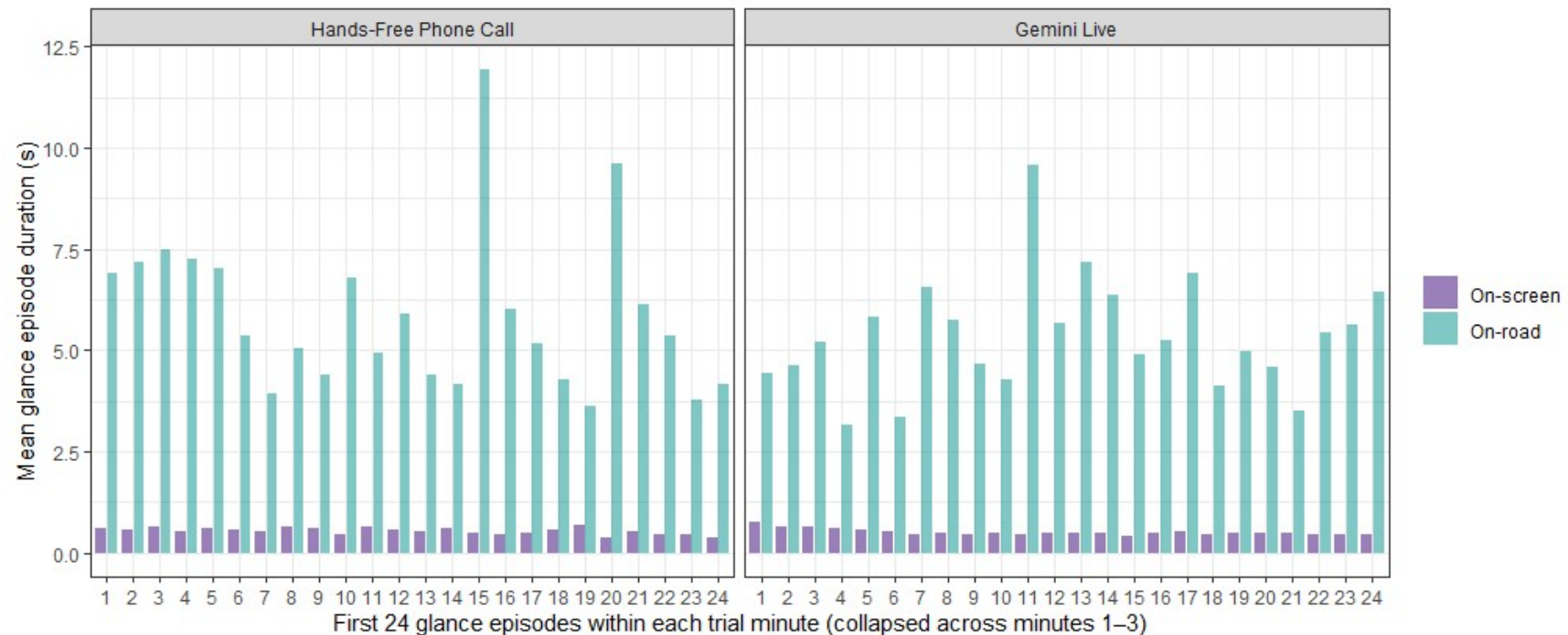
# Glance Episode Durations

- On-road episodes were substantially longer than on-screen episodes
- Gemini Live and hands-free phone conversations showed similar episode durations
- Drivers spent most viewing time monitoring the roadway



# Sequential Structure of Glance Episodes

- Brief glances to the device were embedded within longer roadway-monitoring periods
- Episode patterns remained stable across the interaction
- Sequence structure was similar between interaction types



# Conclusions

---

- Extended conversational AI produced visual attention patterns comparable to conventional hands-free phone conversations
- Drivers maintained sustained roadway monitoring throughout both interaction types
- Glance-episode analysis provides a useful framework for evaluating emerging voice-based in-vehicle technologies

# Thank You

---

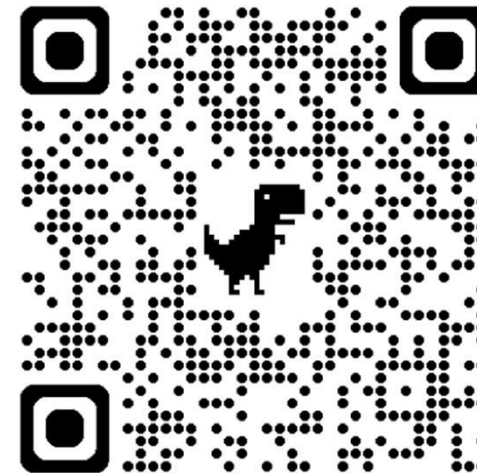
- Contact Info:

- Allegra Ayala, PhD
- Senior Scientist
- aayala@exponent.com

- Contributions from:

- Chris Monk, PhD (Exponent)
- Christine S.P. Yu, PhD (Exponent)
- Ellie Robbins, PhD (Exponent)
- Greg Fitch, PhD (formerly Google)
- Dara Gruber, PhD (Google)

(Preprint version under revision for journal acceptance)



Monk, C., Ayala, A., Yu, C.S.P, Fitch, G.M., & Gruber, D. (2026). Visual and cognitive demands of a large language model-powered in-vehicle conversational agent. arXiv preprint arXiv:2601.15034.