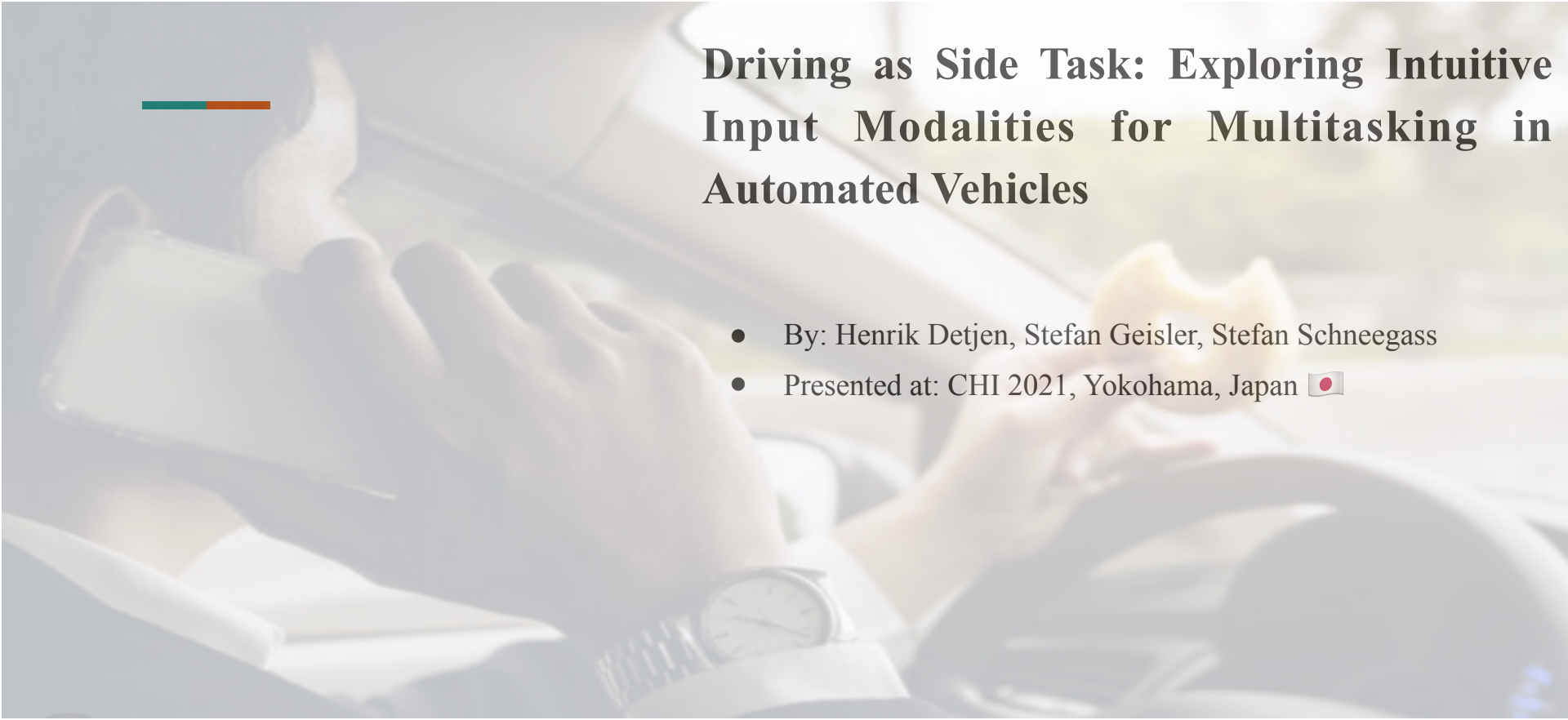


Driving as Side Task: Exploring Intuitive Input Modalities for Multitasking in Automated Vehicles

- By: Henrik Detjen, Stefan Geisler, Stefan Schneegass

- Presented at: CHI 2021, Yokohama, Japan 🇯🇵

Background: Automated Driving & NDRAs

SAE Level 3–5 Automation

Shifts most driving tasks from human to vehicle.

Passive Passengers

Humans become passive during long autonomous phases.

Non-Driving-Related Activities (NDRAs)

Passengers engage in smartphone use, eating, music, conversation.

Occasional Interventions

Vehicles still require user input (e.g., lane change, parking).

L3	Supervisor	Light NDRA
L4	Passenger	Frequent NDRA
L5	Full Passenger	Any NDRA

NDRA–DRA Resource Conflict

- NDRAs consume manual, visual, auditory, and cognitive resources.
- Driving interventions often rely on the same resources.
- Creates resource conflicts, influencing interaction modality choice.

Smartphone	Manual + Visual	Touch / Gesture
Eating	Manual	Gesture
Music	Auditory	Voice
Conversation	Auditory + Cognitive	Voice
Nothing	None	No conflict

Related Work & Research Gap

- Prior studies identified common NDRAs in automated vehicles.
- Takeover research focused mainly on emergency situations.
- Multimodal interaction work showed benefits of switching.

NDRA studies	Types and frequency of NDRAs	Effect of NDRAs on modality choice
Takeover studies	Emergency handover performance	Non-emergency interventions
Multimodal interaction	Mode efficiency and user preference	NDRA-driven modality selection

Gap: We still know little about how everyday NDRAs affect modality choice during normal automated driving.

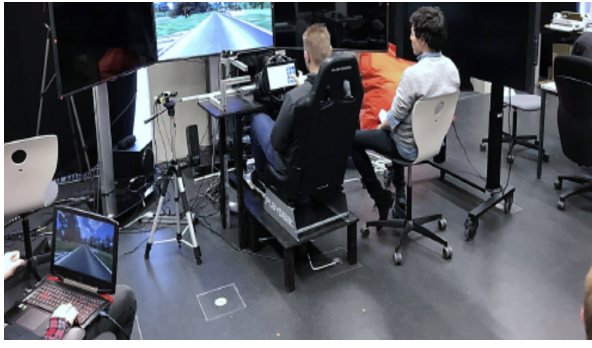
Research Questions & Hypotheses

This study investigates how NDRAs shape users' choices of interaction modality (touch, voice, gesture).

RQ1	Do different NDRAs influence which input modality users choose?
RQ2	Do users avoid modalities that conflict with their current NDRA?
H1	Different NDRAs lead to different modality preferences.
H2	Users avoid modalities that share the same resource demands as their NDRA.

Experiment Setup

- Study on how people control a self-driving car while multitasking.
- 20 computer science students (17 males, 3 females, ages 18–33) participated.
- Used a driving simulator that looked and sounded real.
- Car drove itself, but users could control it via voice, touch, or hand gestures.
- Used Wizard of Oz method a hidden researcher operated the system.
- Driving tasks: start car, join highway, exit, and park.
- Other activities: doing nothing, eating, using a phone, talking, or listening to music.
- After each task, participants filled the NASA-TLX survey to rate mental/physical effort.



Each person tried all five activities in random order. During every activity, they made four small driving actions a total of 20 actions per participant. After each round, they filled out a NASA-TLX form to rate how difficult or stressful the task felt. The entire session took about 30 minutes, and at the end, participants shared which control method (voice, touch, or gesture) they liked the most.

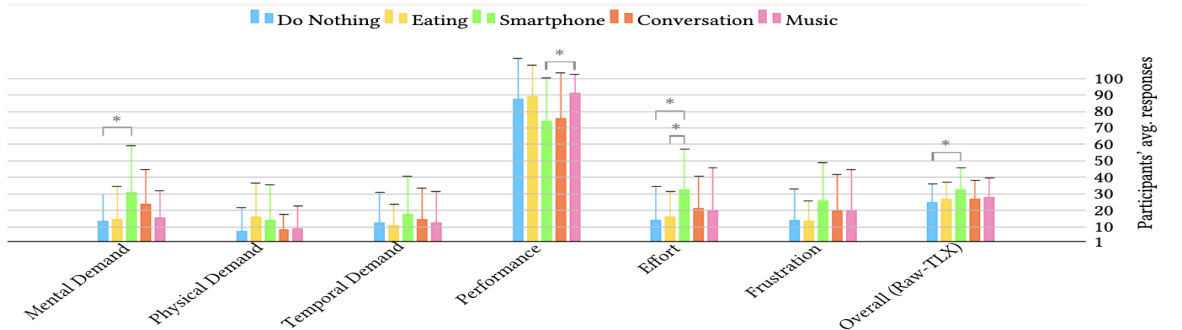
Key Findings

- People's choices regarding control (voice, touch, gesture) varied according to their activity for instance, they did not use voice as much while listening to music and also used fewer gestures when using a smartphone.
- Touch controls were the most popular with 50% since they were perceived as easy and familiar.
- Smartphone usage resulted in the highest mental workload and stress.
- Participants primarily used one or two control types rather than frequent switching.
- All control methods received a positive rating, although there was a slight preference for voice.
- The study suggests the development of multimodal, adaptive interfaces that vary according to users activities.

	Interaction Frequency *n (%)*				**Interaction Variability** *mean, sd*	
Activity	Voice	Touch	Mid-Air Gesture	∑	Flexibility	Change Ratio
Do Nothing	19 (24)	35 (44)	26 (32)	80 (100)	.25 ± .3	.23 ± .31
Eating	13 (16)	39 (49)	28 (35)	80 (100)	.25 ± .30	.22 ± .27
Smartphone*	19 (24)	45* (56)	16* (20)	80 (100)	.28 ± .30	.25 ± .30
Conversation	22 (27)	39 (49)	19 (24)	80 (100)	.18 ± .34	.15 ± .31
Music*	10* (12)	43 (54)	27 (34)	80 (100)	.23 ± .34	.23 ± .36
Overall	83 (21)	201 (50)	116 (29)	400 (100)	.24 ± .32	.22 ± .31

Results

- Touch was the most used control method (50%), followed by gesture (29%) and voice (21%).
- Smartphone use created the highest mental demand and overall workload.
- People avoided using controls that conflicted with their main task (e.g., less voice during music, fewer gestures while talking).
- Participants mostly stuck to one or two control types instead of switching often.
- All input methods were rated positively, with a slight preference for voice control.
- Results highlight the need for multimodal and context-aware interfaces in future automated vehicles.



Why It Matters

- The significance of this outcome lies in the fact that it reveals the influence of different daily activities on the way people communicate with self-driving cars.
- An individual's mental focus and the attention they have available can vary according to the activity they are engaged in, such as using a phone or eating, which in turn determines the control methods (touch, voice, or gesture) that they can use safely and comfortably.
- The high mental workload during smartphone usage suggests that people might have difficulties in paying attention to driving tasks when needed.
- Through the understanding of these effects, the designers can create vehicle interfaces that are safer, easier, and more adaptive, thus, for example, the interface could suggest voice control when hands are busy.
- In conclusion, these revelations are very important for human-vehicle interaction improvement, thus, making autonomous cars not only smart but also user-friendly and safe in the real world multitasking situations.

Limitation & Future Work

Limitations:

- The study used a Wizard of Oz setup, where a hidden researcher controlled the system. Real voice and gesture systems may have higher error rates, so results could differ in real use.
- The sample size was small (only 20 participants) and mostly young male computer science students, limiting how well the findings apply to the general population.
- The study was done in a simulator, not a real car, so it may not fully reflect real-world distractions and environments.

Future Work:

- Test with larger and more diverse participants (different ages, genders, and backgrounds) for broader results.
- Use real voice and gesture systems to evaluate real-world performance and error handling.
- Explore adaptive interfaces that detect what users are doing (e.g., eating, using a phone) and automatically suggest the best interaction mode.
- Study long-term behavior and how people adapt to multimodal systems over time.

Conclusion

- As cars become more autonomous, driving becomes a side task, and people engage in other activities like eating, talking, or using smartphones.
- These non-driving activities affect how people interact with the car.
- Users avoid using controls that interfere with their current activity (e.g., less voice while listening to music, fewer gestures while using a phone).
- Touch controls were used the most, as they felt simple and familiar.
- The findings emphasize the need for multimodal, adaptive interfaces that adjust based on user activity.
- Such systems can make automated vehicles safer, easier to use, and more comfortable for multitasking passengers.

Discussion Questions

- How can future self-driving cars detect what activity (like eating or using a phone) the passenger is doing and automatically choose the most suitable control method?
- Do you think people will trust and feel comfortable using voice or gesture controls more than touch in real self-driving cars? Why or why not?
- What safety challenges might arise if users are too focused on non-driving activities and fail to respond when the car needs human intervention?


THANK YOU
Stay safe. Drive safely.