Smart  Research for In-vehicle Voice Interaction

Auk Kim
kimauk@gmail.com
https://www.kimauk.com/

Papers (*IMWUT – formally known as UbiComp; Top HCI conference)

- Kim A, Choi W, Park J, Kim K, Lee U. "Interrupting Drivers for Interactions: Predicting Opportune Moments for In-vehicle Proactive Auditory-verbal Tasks". In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT),* 2(4), 175. ACM (2018 December).
- Kim A, Lee U. "Interruptibility for In-vehicle Dual Tasking: Influence of Voice Task Workload and Adaptive Behaviors". In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT*), ACM (Revision submitted on 2019 December)

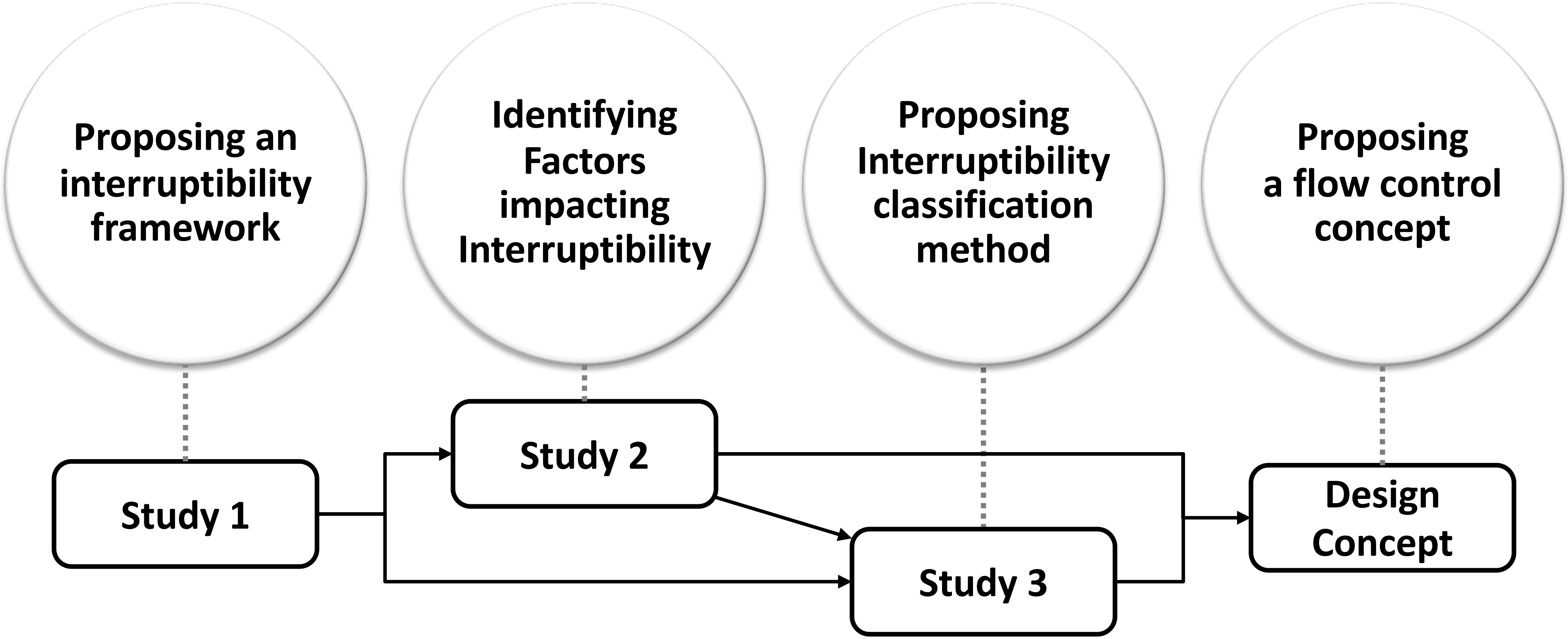
Poster

- Kim A, Choi W, Park J, Kim K, Lee U. "Predicting opportune moments for in-vehicle proactive speech services". In *Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers (UbiComp '19 Poster),* 101-104. ACM (*2019)*.

* Starting with the 2017 edition, UbiComp no longer considers full paper or note submissions. Instead, it will invite for presentation papers published by the *Journal Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*

Research Overview

Developing a method and design that can enable **drivers to safely interact with voice interactions successfully**



- **Study 1:** Developing experimental framework that enable to observe and collecting driver interruptibility in naturalistic driving environments
- **Study 2:** Examining the influences of interruption contexts and driver behaviors on drive interruptibility
- **Study 3:** Developing driver interruptibility prediction model
- **Design Concept:** Towards flow control of driver-vehicle

Media Coverage

Oversee news paper

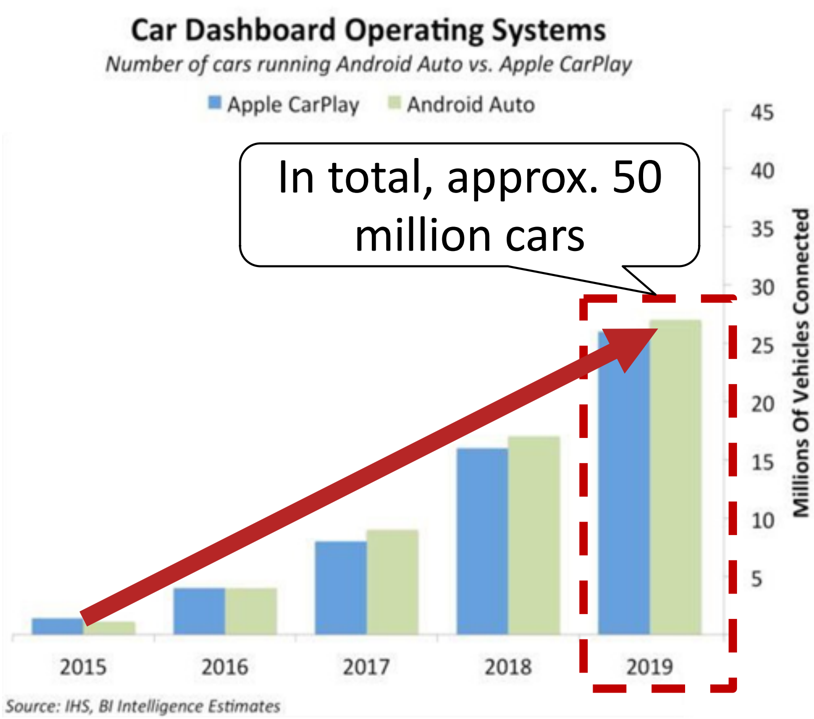
- EurekAlert. AI to determine when to intervene with your driving. November 2019. https://www.eurekalert.org/pub_releases/2019-11/tkai-atd111319.php
- EENews Europe. *"Designing safe context-aware AI assistants for car drivers"* (19.11.13). https://www.eenewseurope.com/news/designing-safe-context-aware-ai-assistants-car-drivers
- ScienceDaily. *"AI to determine when to intervene with your driving"* (19.11.13). https://www.sciencedaily.com/releases/2019/11/191113092604.htm
- The Media Hq. *"AI to establish when to intervene with your driving"* (19.11.13). https://themediahq.com/ai-to-establish-when-to-intervene-with-your-driving/
- Mirage News. "AI to Determine When to Intervene with Your Driving" (19.11.13) https://www.miragenews.com/ai-to-determine-when-to-intervene-with-your-driving/

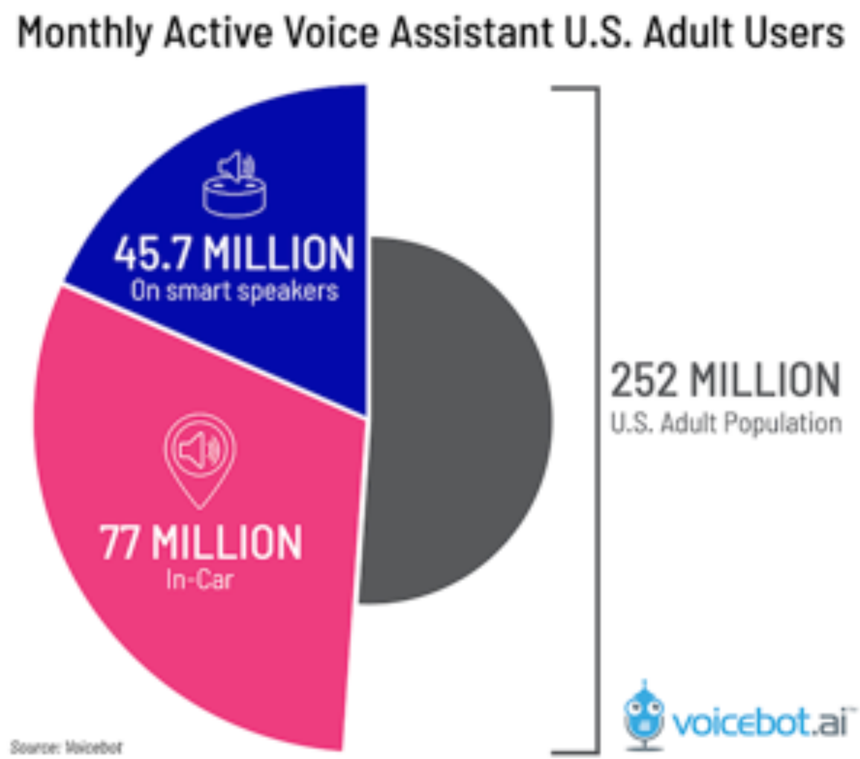
Domestic news paper

- Aju Business Daily (영문판). *"New AI technology developed to predict right time for conversation with driver"* (19.11.11). https://www.ajudaily.com/view/20191111140917422
- 전자신문. *"KAIST, AI 차량 대화 서비스 개발"* (19.11.12). http://www.jeonpa.co.kr/news/articleView.html?idxno=81927
- 한국경제. "말 걸 '타이밍' 아는 인공지능 차량 대화 서비스 개발" (19.11.11). https://www.hankyung.com/society/article/201911118443Y
- 동아사이언스. *"AI로 차량 대화형서비스 과몰입 사고 막는다" (19.11.11).* http://dongascience.donga.com/news.php?idx=32275
- 아시아경제. *"언제 운전자에게 말 걸지 AI가 판단한다"* (19.11.11). https://www.asiae.co.kr/article/2019111111111084272

Introduction: In-vehicle voice interface and agents

In-vehicle voice interface and agents are becoming increasingly popular





Introduction: In-vehicle proactive voice agents

In the future, many in-vehicle agents proactively offering various voice services

TECH

Toyota's Talking Car Wants to Be Your Clingy BFF

AI 'Yui' just needs to track your social-media activity and record what you say

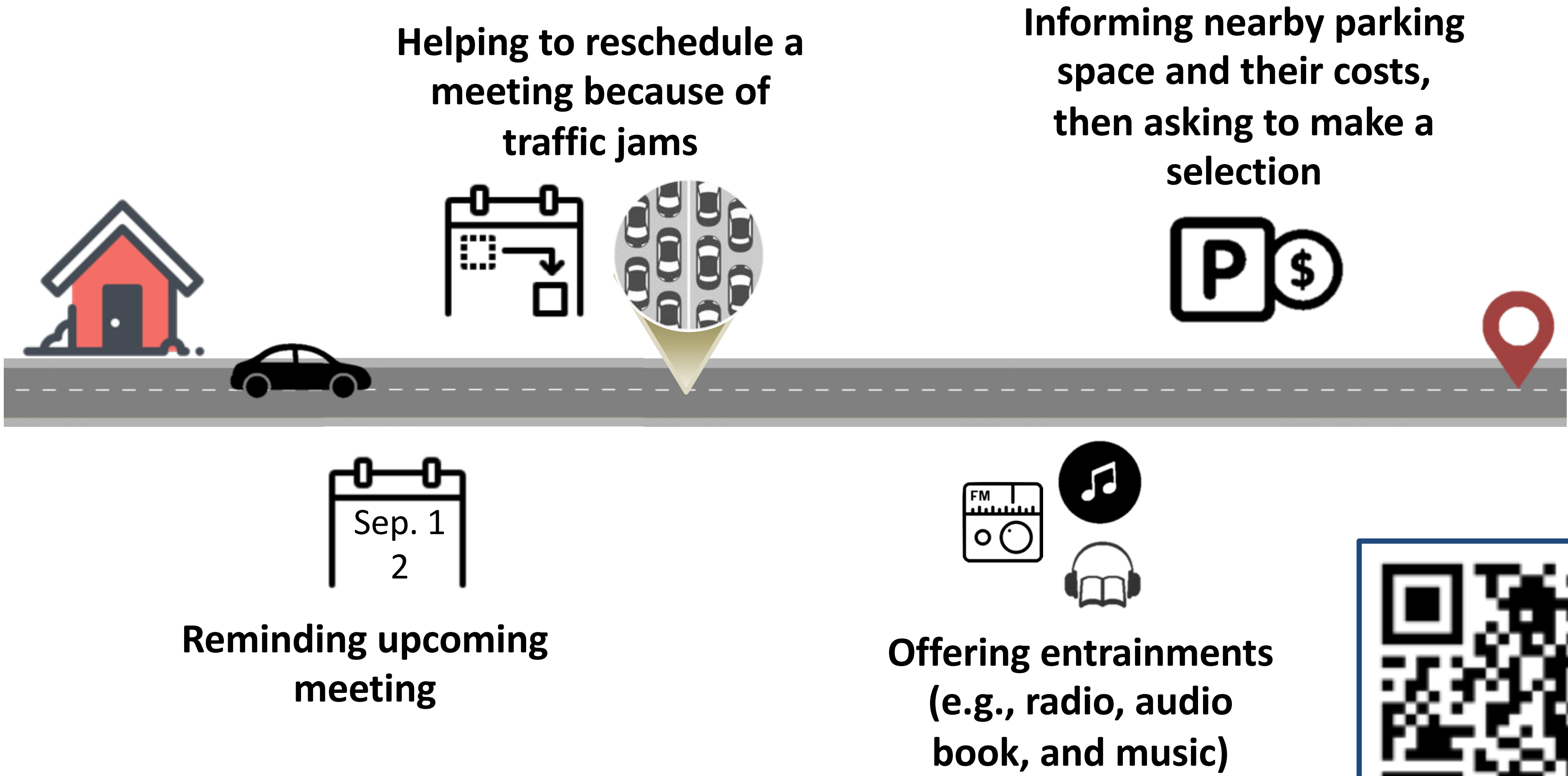
Jaguar Land Rover's Self-learning car

BMW Backs Talking Car 'Intelligent Assistant' Technology

Adrian Bridgwater Senior Contributor
Enterprise & Cloud
I track enterprise software application development & data management.

[1] The Wall Street Journal 2017. Toyota's Talking Car Wants to Be Your Clingy BFF.
[2] Jaguar Land Rover 2019. The world's first self-learning car
[3] Bridgwater 2018. BMW Backs Talking Car 'Intelligent Assistant' Technology

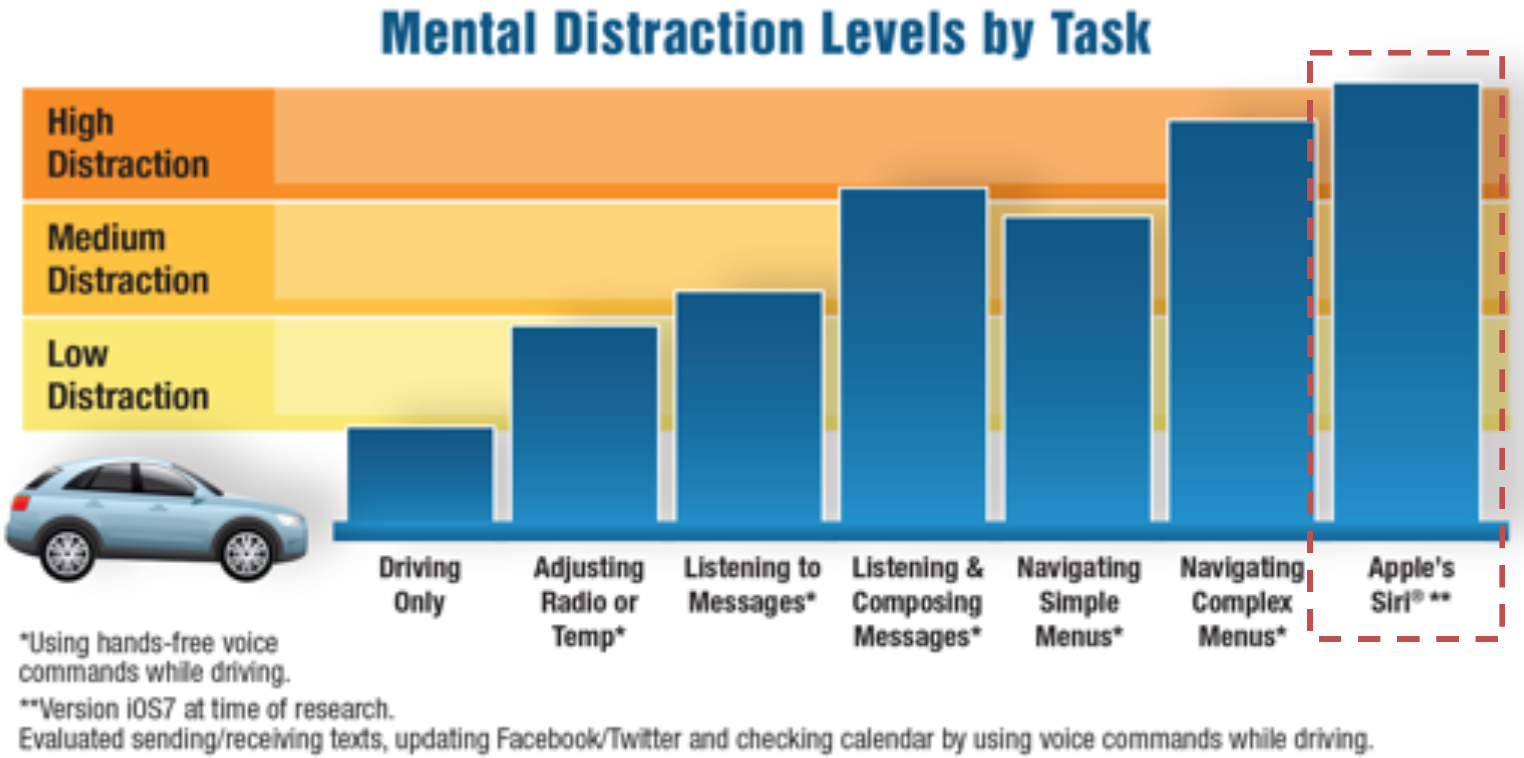
Introduction: Example of in-vehicle proactive voice services scenario



Video Link: **https://youtu.be/kGWFyHSU8aw**

Introduction: Cognitive distraction

Voice interactions (or interruptions) cause cognitive distractions, leading to distracted driving

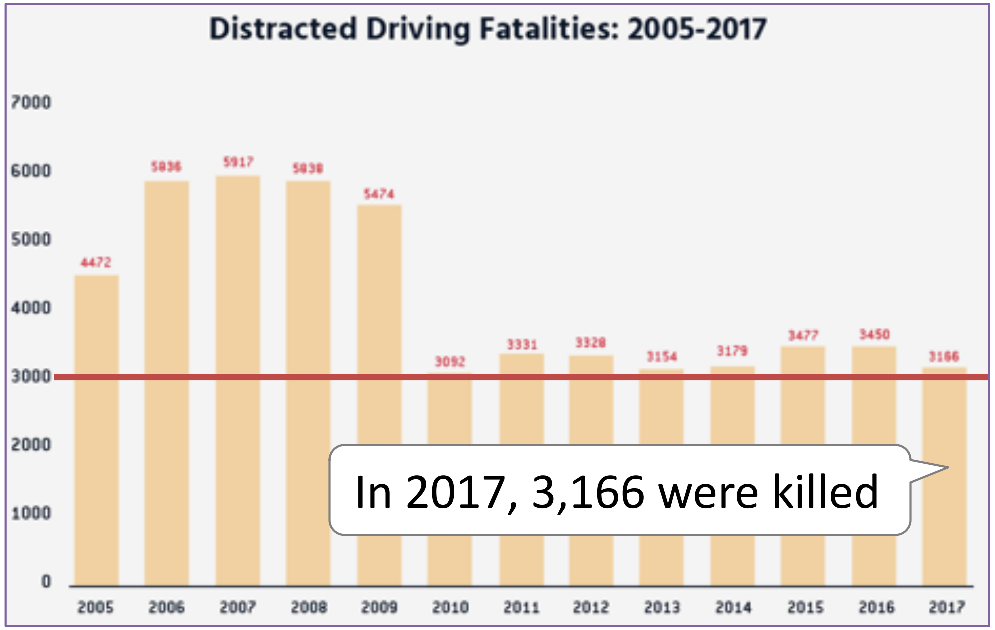


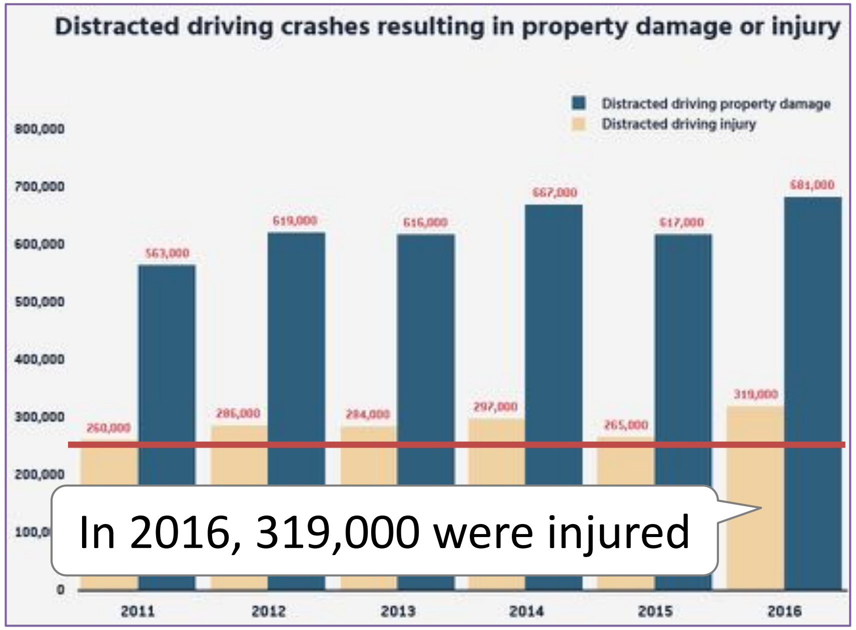
When a driver is cognitively overloaded with driving, the use of speech interfaces can degrade driving performance [Faure 2016]

[1] Strayer 2014. "Mental workload of common voice-based vehicle interactions across six different vehicle systems."
[2] Faure 2016. The effects of driving environment complexity and dual tasking on drivers' mental workload and eye blink behavior

Introduction: Distracted driving is Deadly

In U.S., every year, approx. 3000 people are killed and 319000 are injured





[1] NHTSA 2019, "Traffic Safety Facts Research Notes 2019: Distracted Driving. S. Department of Transportation, Washington, DC: NHTSA; 2019.".

Introduction: Human cognitive biases, errors, and limits

Cognitive biases, errors, and limits encouraging and contributing cognitive distraction

Cognitive biases (e.g. Optimism bias, Risk-tasking behaviors)

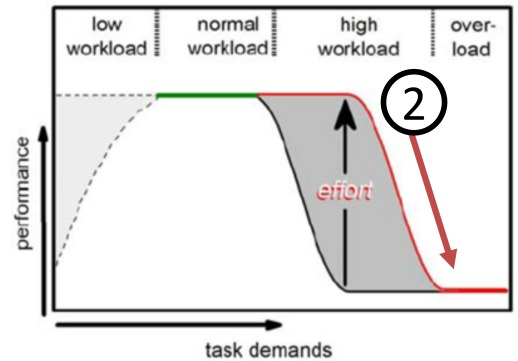
- **Overestimating driving capability, underestimating accident risk**s [DeJoy 1989]
- Neglecting driving because expected rewards (secondary task) outweigh the expected costs [Fuller 1991, Horrey 2009]

Cognitive error (e.g. Contingency traps, Task prioritization error)

- Failing to attend because the **hazards and roadway demands are difficult to perceive** [Fuller 1991]
- Having difficulties in prioritizing one task (e.g., driving) over another task (e.g., secondary task) [Levy 2008]

Cognitive limits (i.e. limited cognitive resources)

- Human has limited cognitive resources. If multitasking demands more than available resources, it cause **performance decrements in either one of, or both driving and secondary tasks** (i.e., dual-tasking interference) [Wickens 2008]



[1] DeJoy 1989. The optimism bias and traffic accident risk perception.
[2] Fuller 1991. Behavior analysis and unsafe driving: Warning-learning trap ahead!
[3] Horrey 2009. Driver-initiated distractions: Examining strategic adaptation for in-vehicle task initiation.
[4] Levy 2008.Task prioritisation in multitasking during driving: opportunity to abort a concurrent task does not insulate braking responses from dual-task slowing.
[5] Wickens 2008. Multiple Resources and Mental Workload.

Introduction: Problem statement

The distracted driving caused by voice interaction is a **real problem** as it tasks significant amounts of deaths and injuries every year.

There needs a **technology or design** that can enable systems to **<u>safely interact</u>** with their drivers via voice interfaces

Introduction: Research overview and key contributions

Developing a method and design that can enable **drivers to safely interact with voice interactions successfully**

- **Study 1:** Developing experimental framework that enable to observe and collecting driver interruptibility in naturalistic driving environments
- **Study 2:** Examining the influences of interruption contexts and driver behaviors on drive interruptibility
- **Study 3:** Developing driver interruptibility prediction model
- **Design Concept:** Towards flow control of driver-vehicle

Study 1: Developing experimental framework

Outlines

① **Key contributions**

② **Framework**

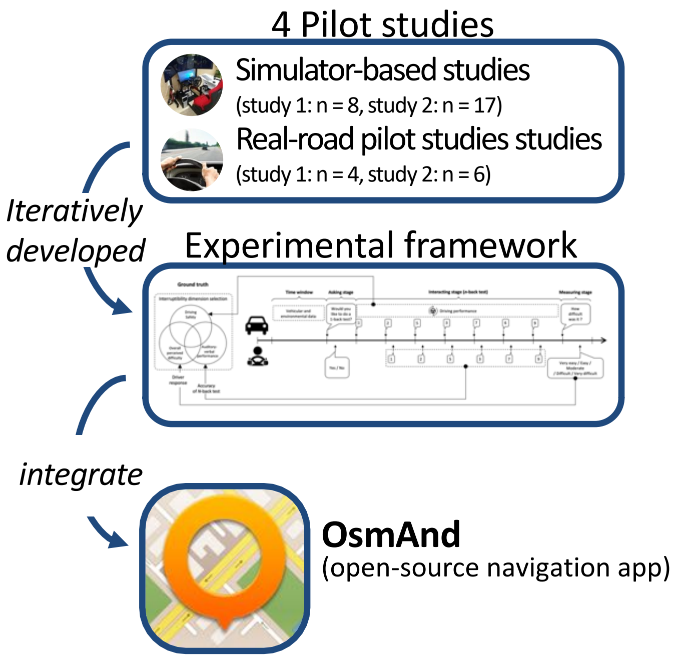
③ **Data-collection Setting**

④ **Driver interruptibility dataset**

Kim A, Choi W, Park J, Kim K, Lee U. Predicting opportune moments for in-vehicle proactive speech services. In *Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers 2019* Sep 9 (pp. 101-104). ACM.

1.1. Key contributions

Driver interruptibility framework

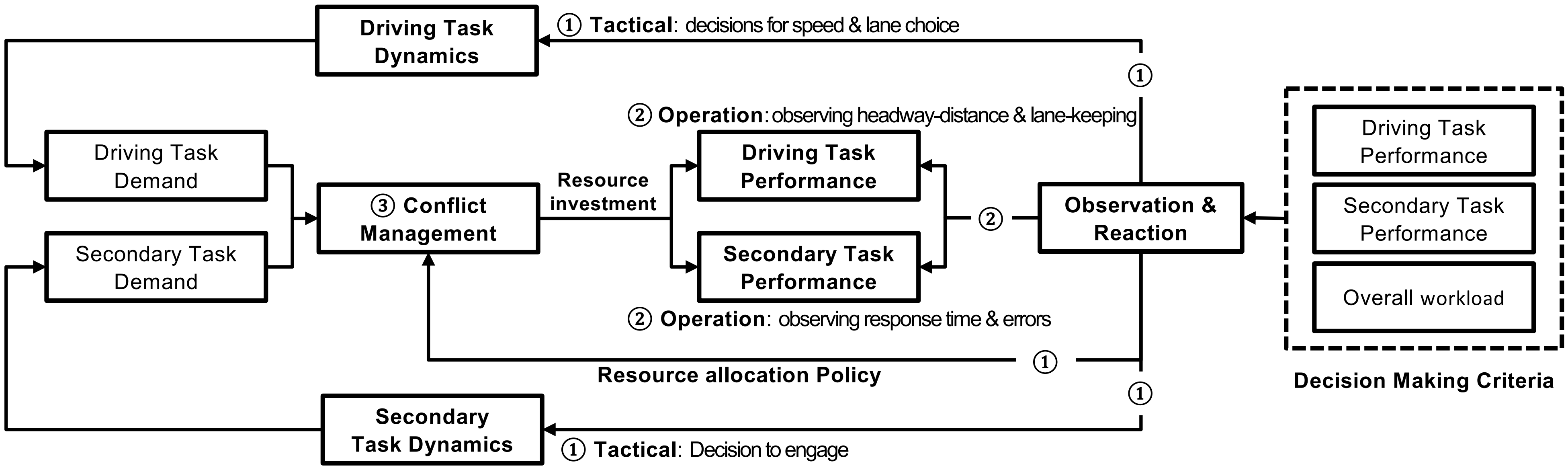


Driver interruptibility dataset

- Iteratively developed a framework (n = 35) that enables to measures driver interruptibility towards voice interactions under various interruption contexts (multiple voice-task level and diverse driving contexts)
- Developed data-collection tools (S/W and H/W) and built driver interruptibility dataset (n = 29) in real-road driving setting.

1.2. Experimental frameworks – Theoretical background

Drivers control driving and secondary tasks according to ① tactical decision, which was made based on ② their observation on the changes in ongoing performance and overall demand (workload) of driving and secondary tasks.



[Control theory-based Hierarchical Driver Model - Lee 2004]

① In tactical level, **making decisions to engage in a secondary task**. When decided to engage, **making decisions for speed, lane choice, and resource allocation policy** based on previous observation results.

② In operation level, according to the tactical-level decision, controlling headway-distance, lane-keeping, and resource investment (resource allocation policy), while **observing the changes in ongoing performances and demands of driving and secondary tasks**,

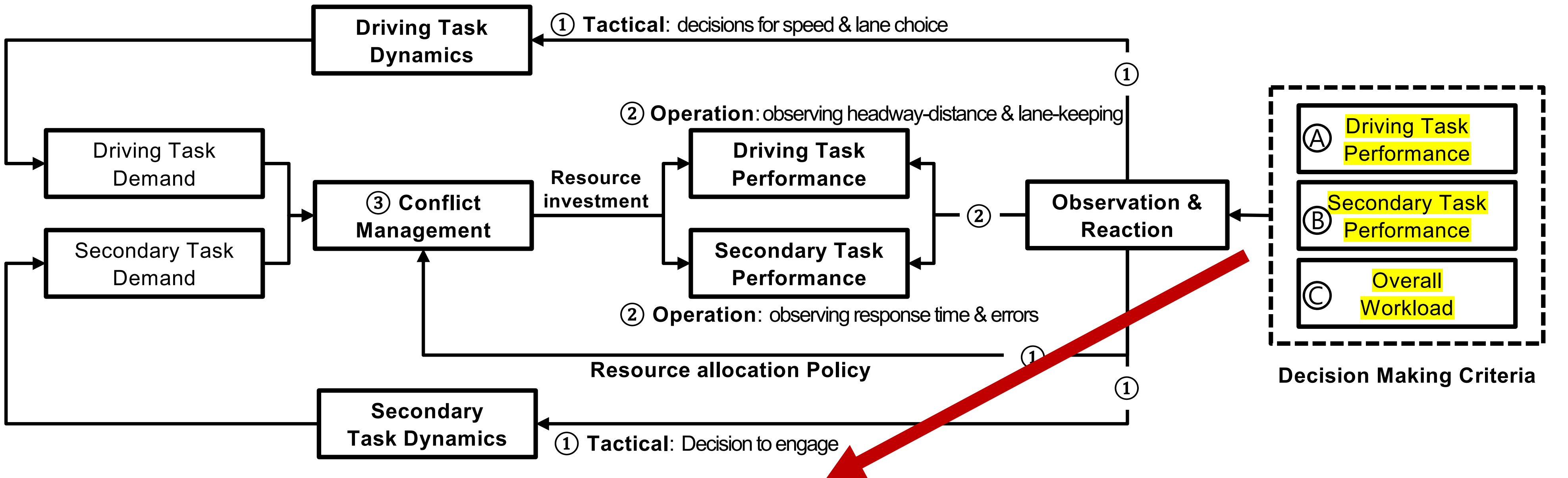
③ (Conflict Management) When **observing any conflicts (e.g., performance decrements & an increase in demand levels), making a new tactical decision**.

[1] Regan 2008. Driver distraction: Theory, effects, and mitigation.
[2] Lee 2004. Preface to the special section on driver distraction.

1.2. Experimental frameworks – Driver *interruptibility dimensions

*Quality of being interruptible by a system

Proposing driver interruptibility dimensions to find opportune moments based on the theoretical grounds (driver model, multiple resource theory)



Driver interruptibility dimensions

A. Driving safety: 'how safely a user drives a vehicle'
B. Auditory-verbal performance: 'how well a user performs an auditory-verbal task'
C. Overall perceived difficulty: Overall perceived difficulty: 'how difficult it is to perform a dual task'

tasks, **making a new tactical decision.**

[1] Regan 2008. Driver distraction: Theory, effects, and mitigation.
[2] Lee 2004. Preface to the special section on driver distraction.

1.2. Experimental frameworks - Overview

Secondary task

Triggering method

Metrics

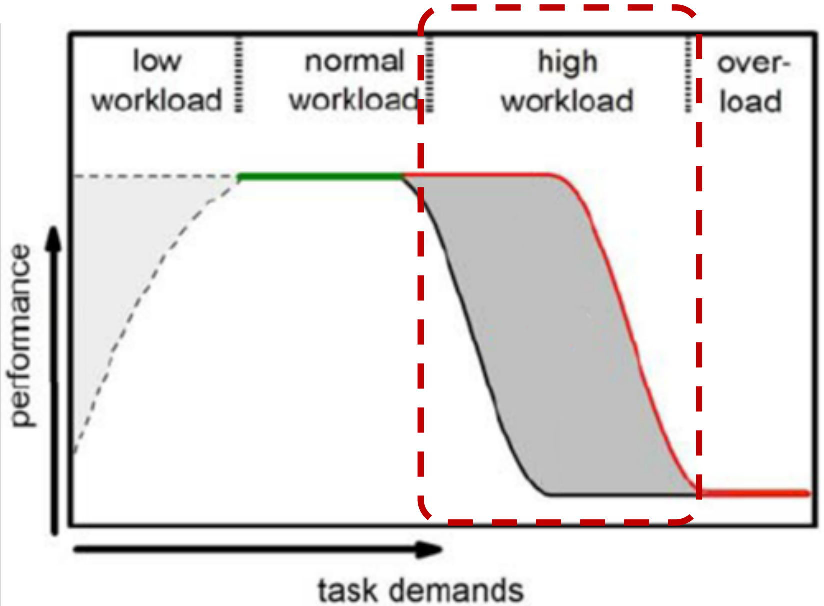
Enable to observe driver interruptibility under diverse interruption contexts

Enable to measure driver interruptibility from various perspectives

1.2. Experimental frameworks - Secondary task

Voice-task sets with variable cognitive demands are required

- Alexa has more than 80 skill (or service) sets that may significantly vary in their cognitive demand levels [Amazone 2019]



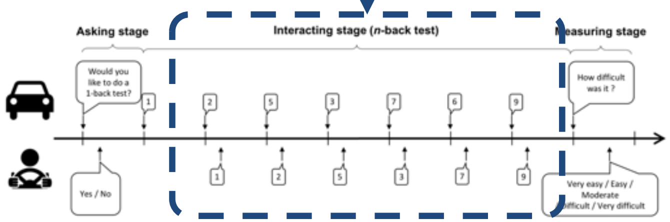
Performance is more likely decreased as demand levels (workload) of secondary task increases [Wickens 2008]

[1] Amazone 2019. https://www.amazon.com/alexa-skills/b?ie=UTF8&node=13727921011
[2] Wickens 2008. Multiple Resources and Mental Workload.

1.2. Experimental frameworks - Secondary task

N-back tasks[Mehler 2019] as an embedded task in secondary task

		Time							Description
		t	t + 2.25s	t + 5s	t + 7.25s	t + 10s	t + 12.25s	t + 15s	
Stimulus		1	2	5	6	7	6	9	Presenting seven randomly selected numbers (0 to 9) at 2.25s intervals
Driver response	***0*-back (very mild demand)**		1	2	5	6	7	6	Repeating each number as it is presented
	***1*-back (moderate demand)**			1	2	5	6	7	Repeating the number one item back in the sequence
	***2*-back (high demand)**				1	2	5	6	Repeating the number two item back in the sequence



Secondary task

- Having a <u>similar cognitive engagement</u> (e.g., auditory attention and memory components) as when a driver engages in an externally paced verbal task such as interacting with the voice interface of in-vehicle devices or answering a cell phone.
- <u>Inducing consistently structured levels of cognitive demand</u>, because we can vary the number of delayed digits and the length of interactions.
- Being widely utilized as an distractive voice-secondary task since 1993 [Zeitlin] <u>in numerous real-road driving studies,</u> including ISO-associated standards research [Bruyas 2015], and prior works carried out by the U.S. National Highway and Transportation Safety Administration (e.g., "Developing a Test to Measure Distraction Potential of In-Vehicle information System Tasks in Production Vehicles" [NHTSA 2011]).

[1] Mehler 2019. Mit agelab delayed digit recall task (n-back).
[2] Zeitlin 1993. Subsidiary task measures of driver mental workload: A long-term field study.
[3] Bruyas 2015. Sensitivity of detection response task (drt) to driving demand and task difficulty.
[4] NHTSA 2011. Developing a test to measure distraction potential of in-vehicle information system tasks in production vehicles.

1.2. Experimental frameworks - Secondary task

Various voice-interaction scenarios in simulator-based studies,
more realistic than n-back tasks, but may not ensure ‘*driving safety* in real-road driving setting’

Phone usage [Strayer 2015]

Music selection [Mehler 2014, Strayer 2015]

Email manipulation [Mehler 2014]

Navigation control [Mehler 2014, Strayer 2015]

Simple factoid Q&A tasks [Rajab 2016]
(e.g., true/false answering of factual questions)

Topic-based free conversation [Large 2016]

- Simulator-based studies have considered various voice-interaction scenarios.
- Although these kinds of scenario-based tasks are more realistic than n-back tasks, designing and developing standardized scenarios with variable cognitive demands is very challenging as that requires considerable time and resources for assuring ***consistency*** and ***reliability***, as well as ***safety validation*** in real-world driving.

[1] Strayer 2015. Measuring cognitive distraction in the automobile iii: A comparison of ten 2015 in-vehicle information systems.
[2] Mehler2014. Further evaluation of the effects of a production level “voice-command” interface on driver behavior
[3] Rajan 2016. Task load estimation and mediation using psycho- physiological measures.
[4] Large 2016. Assessing cognitive demand during natural language interactions with a digital driving assistant.

1.2. Experimental frameworks - Secondary task

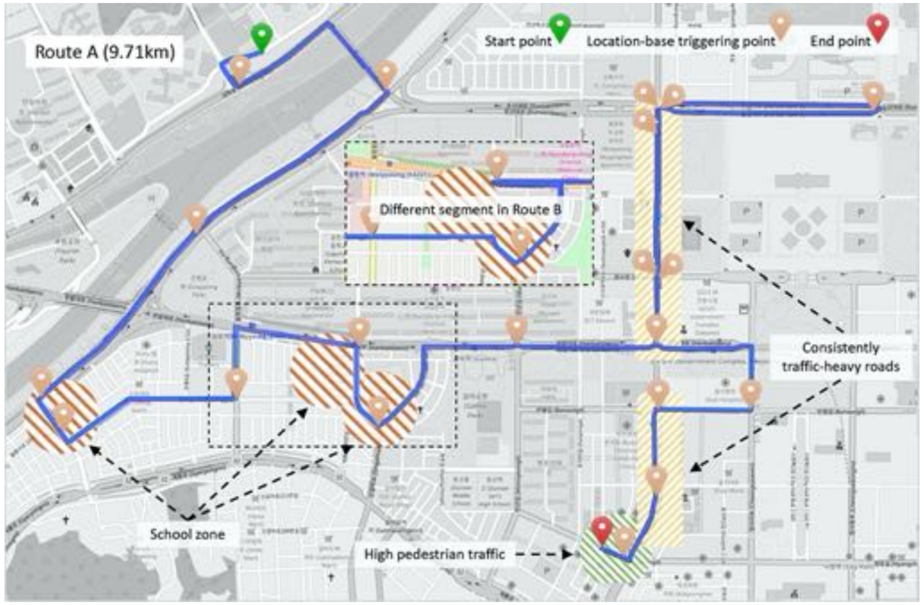
Inducing three varying cognitive levels of voice tasks (n-back tests)

1.2. Experimental frameworks - Triggering method

Hybrid approach of location-based and random-interval triggering methods

Random-interval triggering:
presenting additional tasks at random intervals between 30 and 90 seconds

Location-based triggering:
presenting tasks at for predetermined locations. For the data collection, there were 40 predetermined locations (20 for each route)

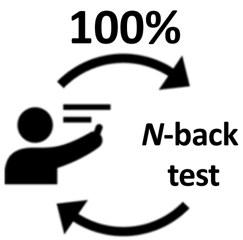


- Route A in driving course. Route B is slightly different from Route A (see dotted box)-

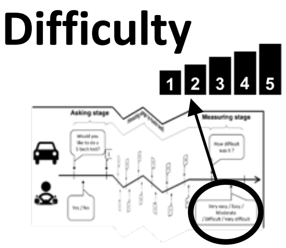
1.2. Experimental frameworks - Metrics

Metrics allow to measure driver interruptibility from various perspectives

Driving safety: *'how safely a user drives a vehicle'*
Interruptible when driving performances (steering wheel reversal rate) during concurrently multitasking of driving and secondary tasks is higher than or equal to that of baseline (i.e., driving alone)



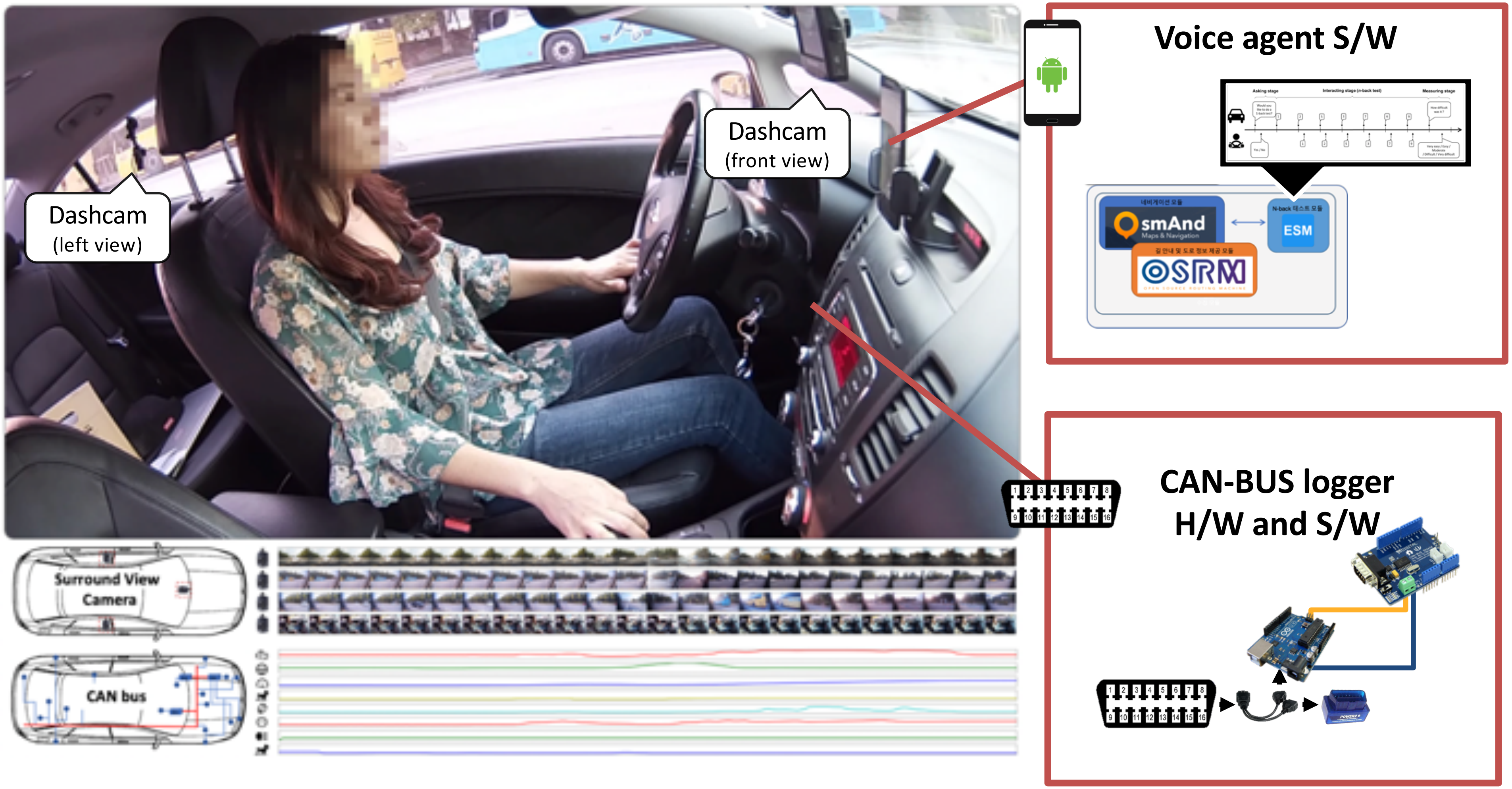
Auditory-verbal performance: *'how well a user performs an auditory-verbal task'*
Interruptible when correctly answering for a given n-back test



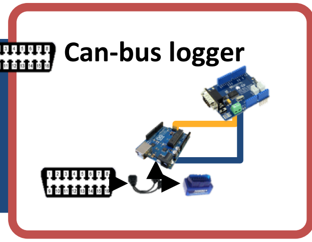
Overall perceived difficulty: *'how difficult it is to perform a dual task'*
Interruptible when the difficulty rating in measuring stage was lower (less than or equal to 0.5)

1.3. Data-collection setting – Equipment

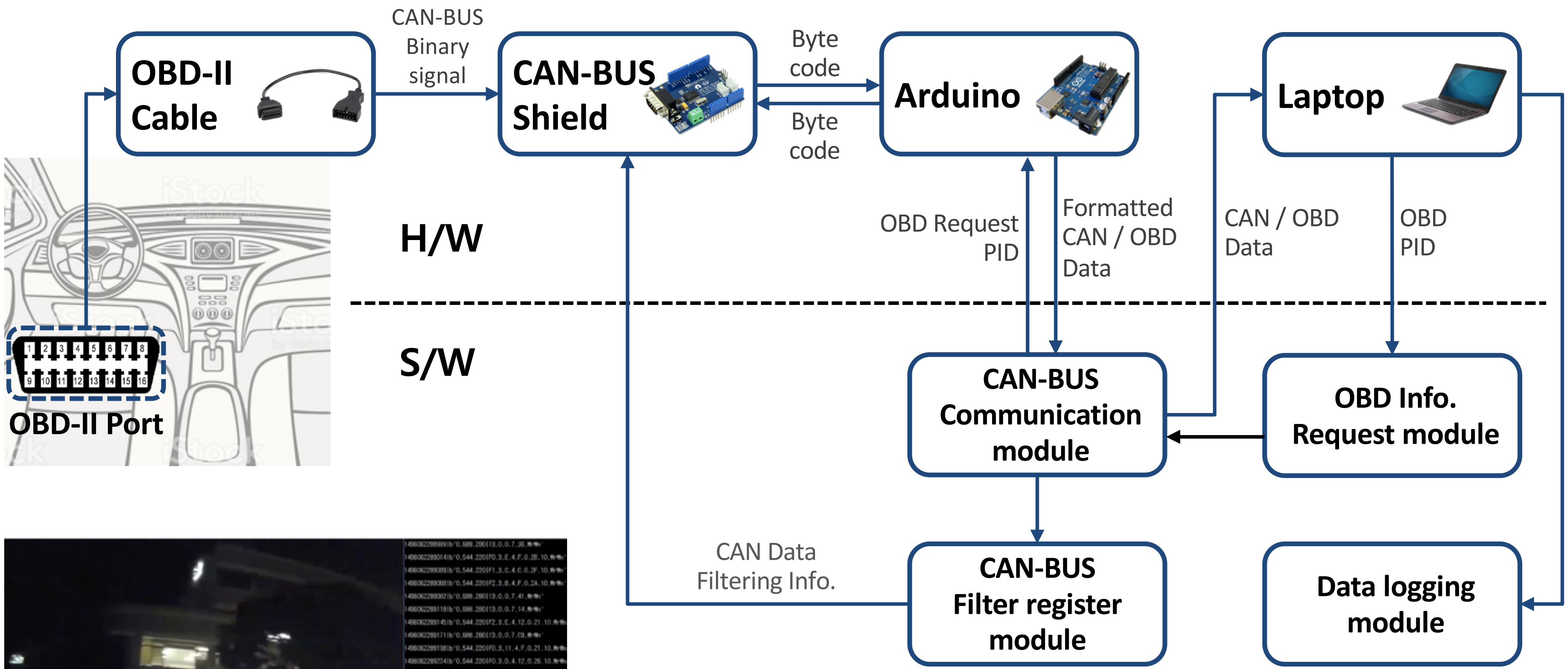
Collecting vehicular and environmental data and driver-agent interaction data



1.3. Data-collection setting – Equipment

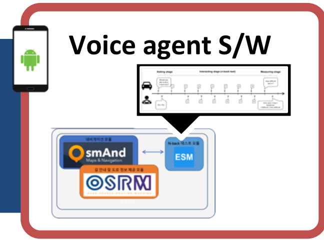


Developing CAN-BUS Logger (Arduino, C, python) that collects vehicle data (e.g., speed, RPM, etc.)

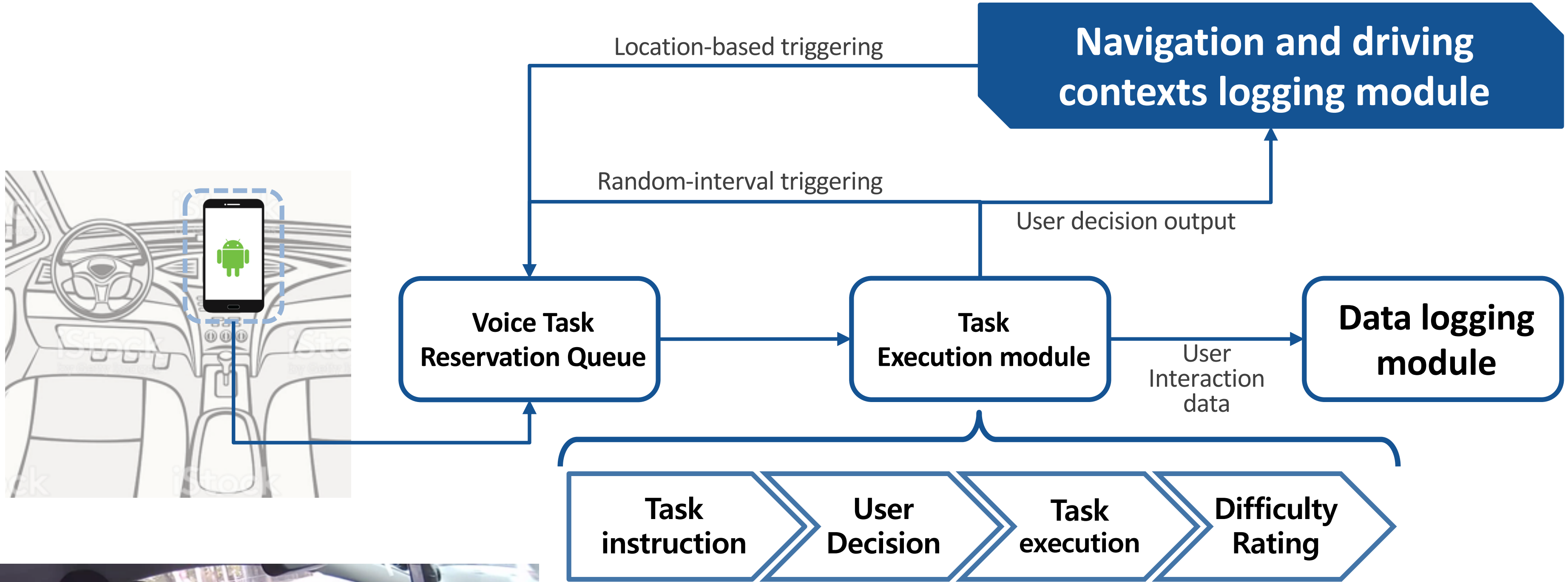


Video Link: https://youtu.be/9jLMXrlSX2M

1.3. Data-collection setting – Equipment

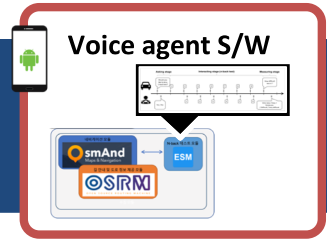


Developing a voice-agent system (Android, JSP)

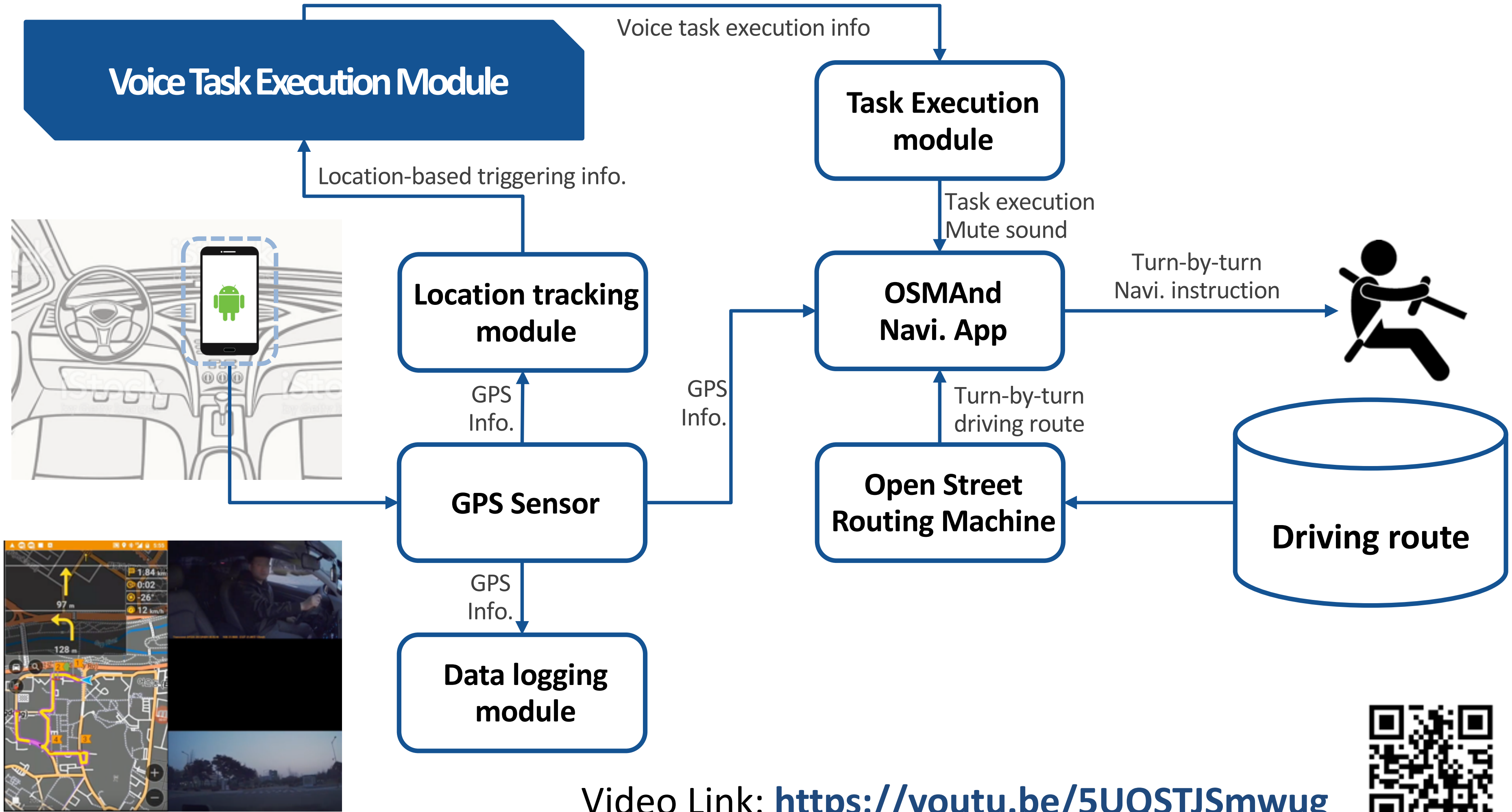


Video Link: **https://youtu.be/nmAI_dh6tUU**

1.3. Data-collection setting – Equipment



Developing customized navigation App (Android) & Server (JSP) for task triggering



Video Link: **https://youtu.be/5UQSTJSmwug**

1.3. Data-collection setting – Collected data types

Vehicle data

- Steering wheel angle
- Throttle position
- Engine position
- Accel. pedal position
- RPM
- Speed
- Turn indicator
- Brake pressure
- Accelerometer

Driving environmental data

- Type of maneuver
- Number of vehicles in the front, left-lane, right-lane of the vehicle
- Distant to adjacent vehicles in the front, left-lane, right-lane of the vehicle

Auditory-verbal interactions

Auditory-verbal interaction data

- *N*-back type
- Overall perc. Difficulty
- Test accuracy
- Time of each stage
- Location
- Audio
- Driver motion (dashcam)
- location

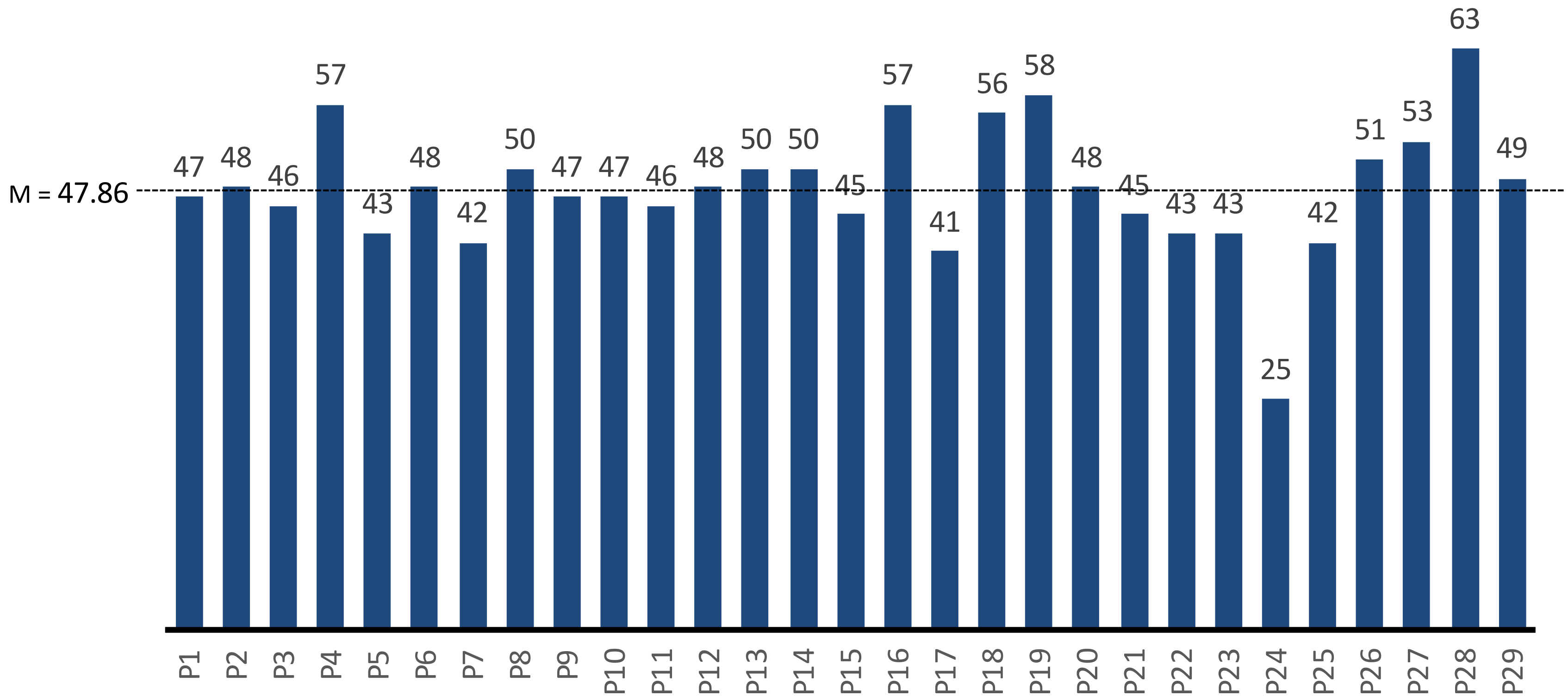
Survey and In-depth interview data

- Demographics
- General vehicle usage
- Driving
- Driving behavior questionnaire
- Driving skill inventory
- Interview results
- Road familiarity
- Reason for test failure

1.4. Data interruptibility dataset (n = 29)

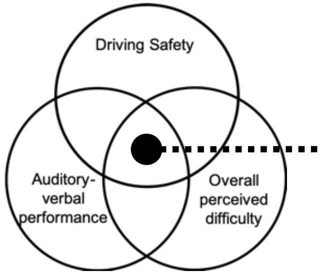
3,480 minutes of a real-world driving dataset with 1388 cases

Average 47.86 (SD = 6.83) cases for a driver



1.4. Data interruptibility dataset – Interruptibility distribution

939 interruptible cases (68%) vs. 449 uninterruptible cases (32%)



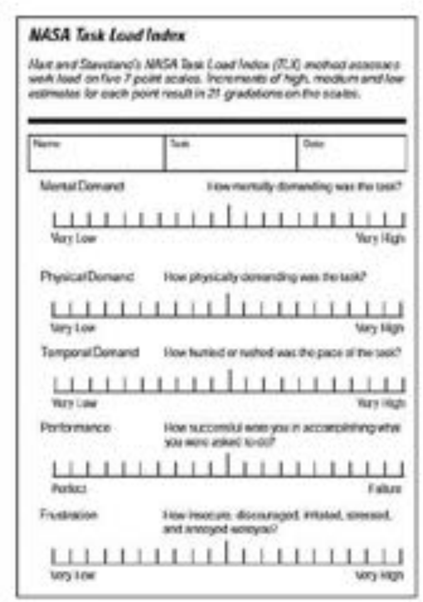
Labeling a secondary task as interruptible if all three *dimensions* are interruptible

Mean (SD)		Overall	Cognitive level of secondary task		
			Very mild (0-back)	**Moderate (1-back)**	**Very high (2-back)**
Overall		0.68 (0.18)	0.88 (0.10)	0.67 (0.25)	0.51 (0.28)
Maneuver type	**Stop**	0.68 (0.19)	0.88 (0.11)	0.65 (0.25)	0.52 (0.32)
	Straight	0.68 (0.17)	0.88 (0.10)	0.67 (0.24)	0.51 (0.28)
	Turn	0.62 (0.22)	0.82 (0.20)	0.61 (0.36)	0.41 (0.36)
	Lane change	0.60 (0.28)	0.76 (0.36)	0.67 (0.43)	0.25 (0.40)

- Average percentage of interruptible moments for a driver -

1.5. Discussion: Mapping n-back to Real-world voice tasks

Each cognitive level of our secondary tasks (or n-back test) can be used as *a reference task* for investigating the interruptibility of real-world voice interactions.



NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7 point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

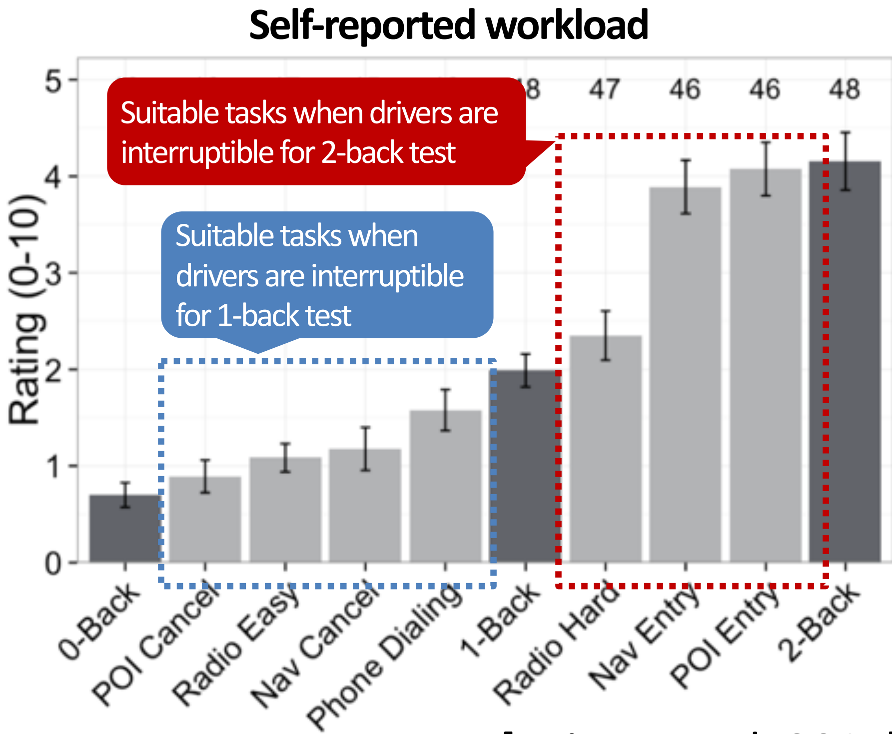
Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Frustration — How insecure, discouraged, irritated, stressed, and annoyed were you?

Very Low — Very High

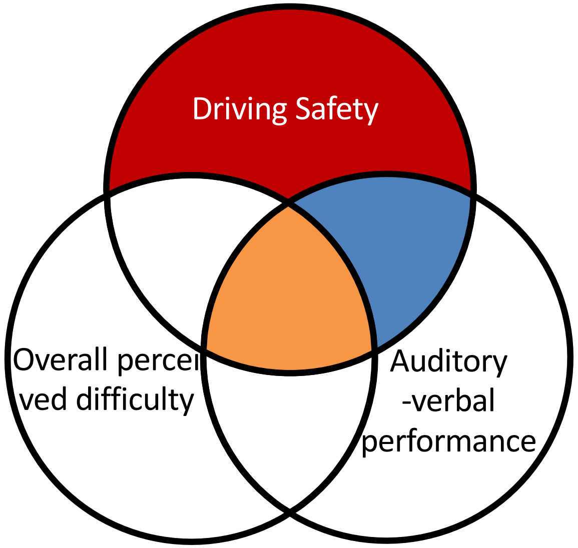


[Reimer et al. 2015]

[1] Reimer 2015. "Phase II Experiment 3 - 2015 Toyota Corolla ”

1.5. Discussion: Selection of interruptibility dimensions

- Possible to use a combination of interruptibility dimensions both disjunctively and conjunctively depending on the purpose of the voice interruptions.
- But, driving safety must be always included as this is always the top priority for any in-vehicle system.



Driving safety (safety)

- appropriate for cases of delivering simple-context information, such as navigation information and weather forecasts.

Safety + Auditory-verbal performance (secondary)

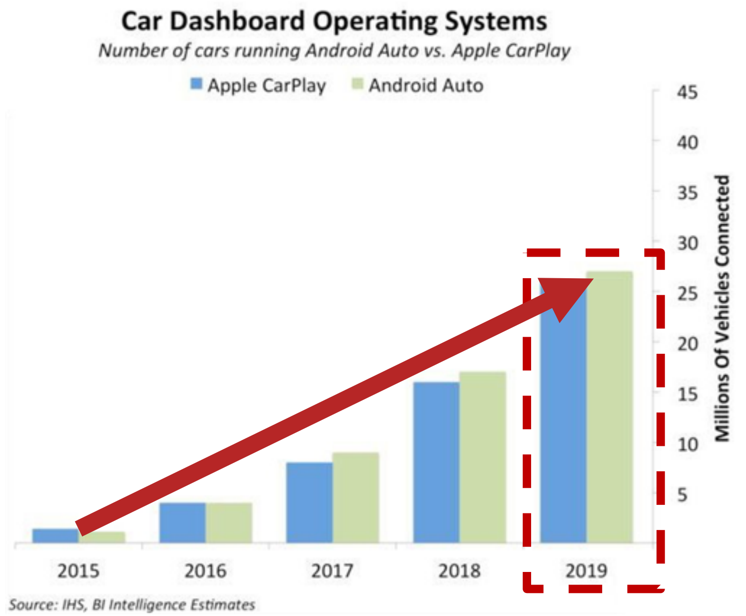
- appropriate for cases of delivering critical information that requires a driver response, such as informing about a low fuel level and asking for permission to reroute to a nearby gas station.

Safety + Secondary + Overall perceived difficulty

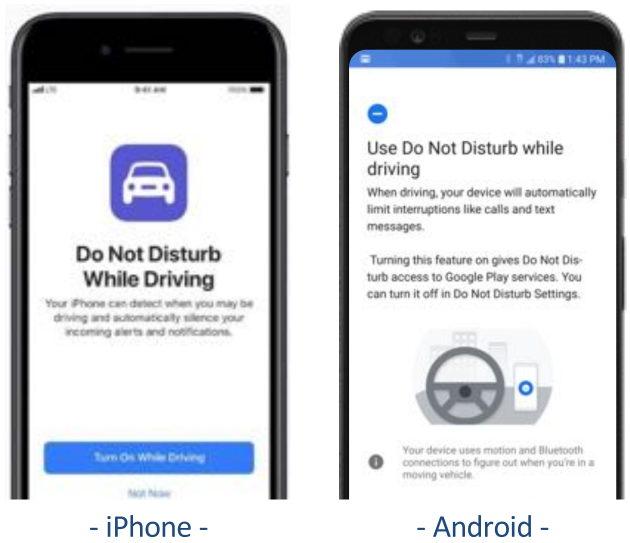
- could be used for voice tasks that involve interactive decision-making (e.g., comparing trade-offs between cost and convenience for choosing parking spots).

1.5. Discussion: Future study directions

Requiring a well-defined voice task set, more realistic than n-back tasks, for research involving voice interaction in real-road driving setting



In recent years, in-vehicle voice interactions have continued to gain popularity, and future interfaces will increasingly rely on the interactions



U.S. NHTSA encourages locking up any interaction tasks that require unreasonable visual-manual distraction

- *N*-back tasks have been widely used in in-vehicle system research involving real-road driving
- It is desirable to compile a well-defined verbal task set for such research where voice agent playing important role.
- In practice, however, it is difficult to design standardized task sets, owing to contextualized responses and multi-turn-taking patterns.
- There should be follow-up studies on these issues, beyond large-scale data collection and validation.

[1] Strayer 2015. Measuring cognitive distraction in the automobile iii: A comparison of ten 2015 in-vehicle information systems.
[2] Mehler2014. Further evaluation of the effects of a production level “voice-command” interface on driver behavior
[3] Rajan 2016. Task load estimation and mediation using psycho- physiological measures.
[4] Large 2016. Assessing cognitive demand during natural language interactions with a digital driving assistant.

Study 2: Examining influences of interruption contexts and driver behaviors on drive interruptibility

Outlines

① Key contributions

② Influences of interruption contexts

③ Influences of driver behavior

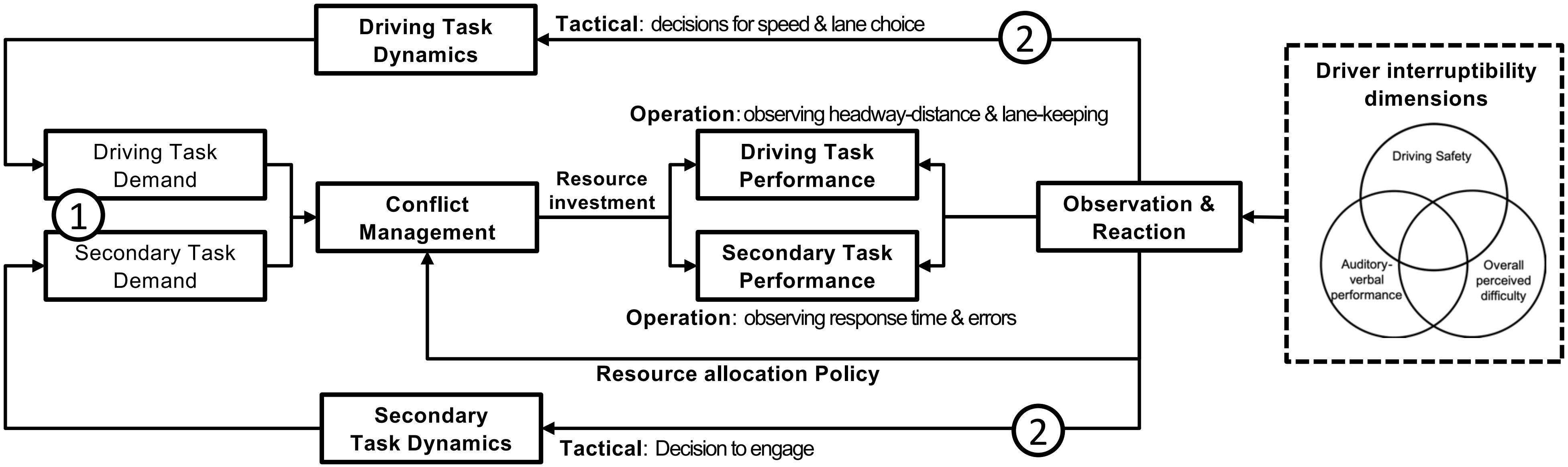
④ Discussion

Kim A, Lee U. "Interruptibility for In-vehicle Dual Tasking: Influence of Voice Task Workload and Adaptive Behaviors". In Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies. 2019 (Revision submitted)

2.1. Key contributions

① **Examining the influences of interruption contexts different interruptibility measures**

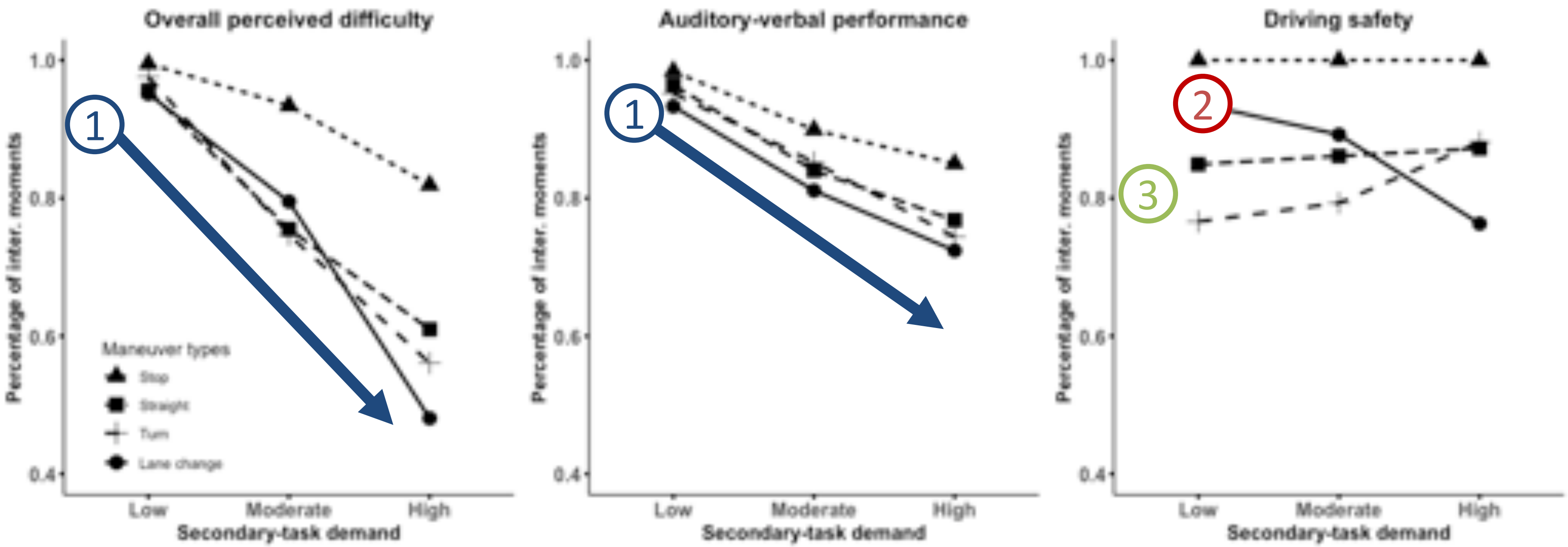
② **Identifying driver behaviors related to interruptibility**



[Control theory-based Hierarchical Driver Model - Lee 2004]

2.2. Influence of interruption contexts on driver interruptibility

Driving demand (maneuver types), as well as secondary-task demand (n-back types) significantly affect driver interruptibility



- Overall, (1) interruptibility decreases as either one or both task demands increases, expect driving safety – (2) for STRAIGHT and TURN, interruptibility does not decreased regardless of an increase in secondary-task demands
- *Prior studies on driver interruptibility have not varied secondary-task demand, while the findings show secondary-task demand significantly affect driver interruptibility

2.3. Influences of driver behavior – Types of driver behaviors

Drivers employed various strategies to maintain their safety and interruptibility

Speed reduction

(n = 23, 79%) reduced the vehicle speed to increase driving safety and interruptibility for the interactions.

Reducing engagement in secondary tasks

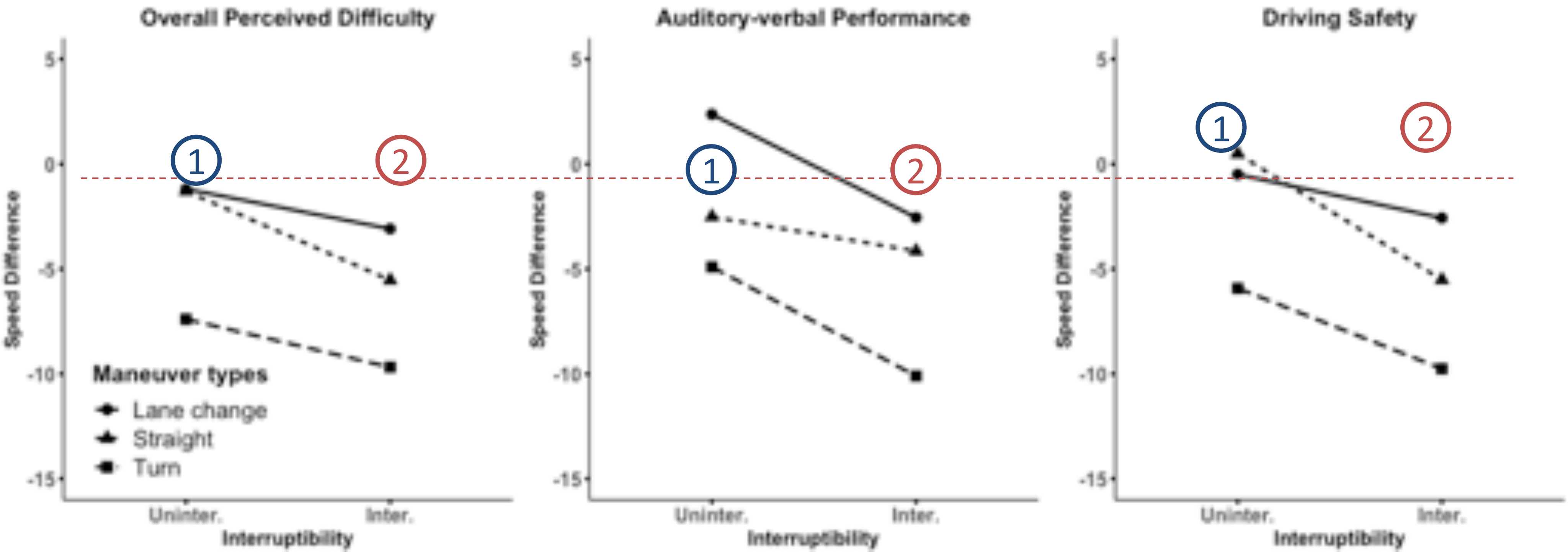
(n = 24, 83%) reduced the attention or ceased to engage in secondary tasks whenever drivers needed to pay more attention to the driving task.

Other adaptive behaviors

- (n = 8, 28%) delayed the vehicle start
- (n = 7, 24%) prepared for interactions by adjusting their driving behaviors in advance (e.g., advanced lane change)

2.3. Influences of driver behavior- Impacts of Speed reduction on interruptibility

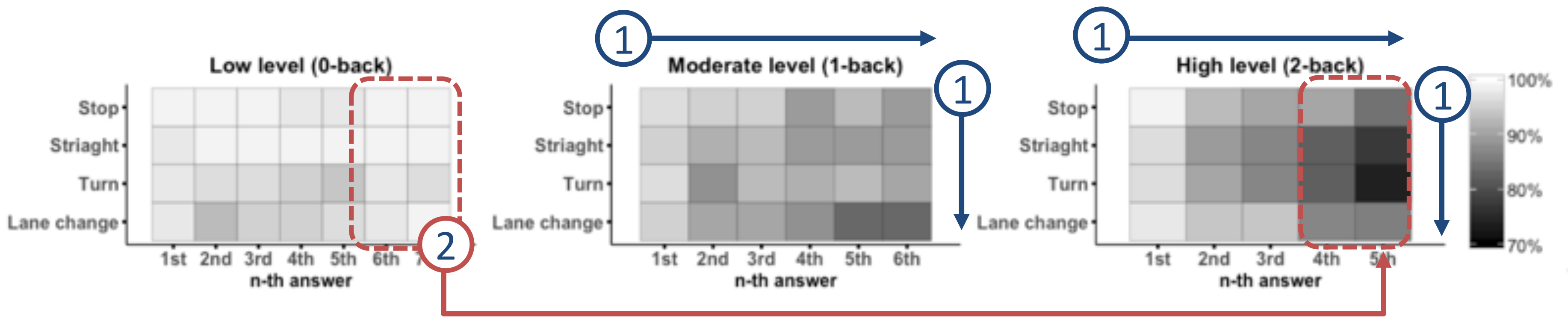
Regardless of maneuver types, speed reduction behavior improves driver interruptibility.



Drivers reduced their speed to a greater extent for TURN than for STRAIGHT and LANE CHANGE, while the speed reduction was greater for interruptible than uninterruptible cases.

The length of secondary tasks affects driver interruptibility

1 2 **Getting darker (more likely to answer incorrectly for a given item)**



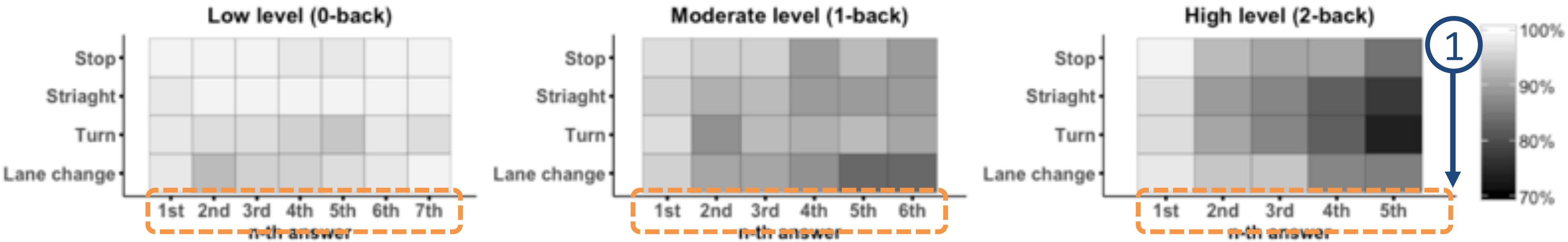
- Heat maps show the percentage of drivers who correctly answered each n-th item in the n-back task across maneuver and n-back types -

- Drivers are more likely answer incorrectly as (1) the item order (n-th) increases, and decreases further as (2) the demands of n-back type increase
- *None of the prior studies on driver interruptibility have either examined the impact of the length of a secondary task on interruptibility or considered the length in interruptibility

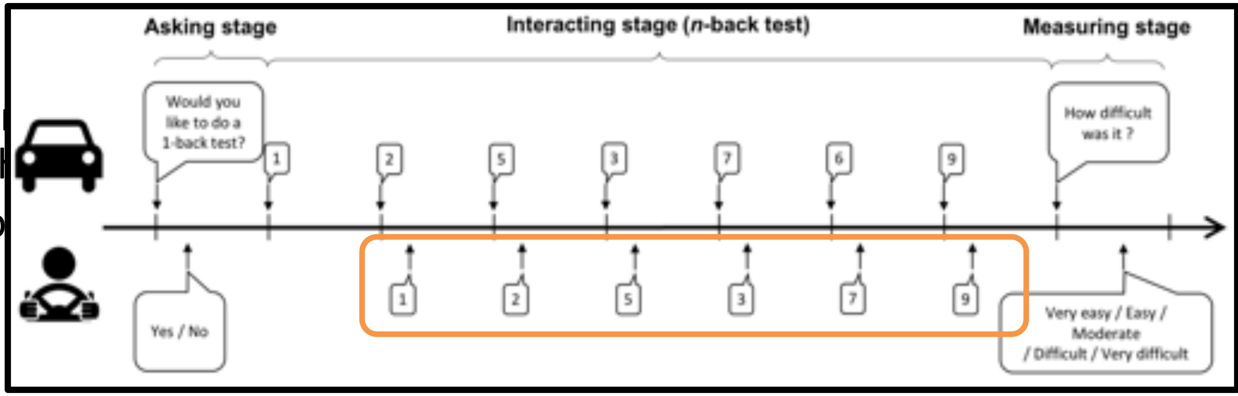
2.3. Influences of driver behavior - Impacts of reducing engagement in secondary tasks on interruptibility

The length of secondary tasks affects driver interruptibility

Getting darker (more likely to answer incorrectly for a given item)



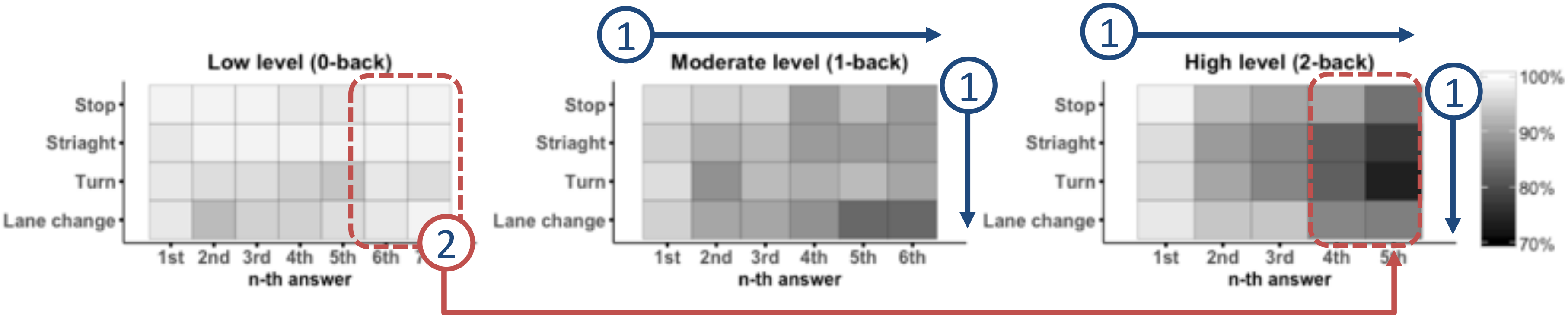
- Heat maps show the percentage of drivers who correctly answered each n-th item in the n-back task across maneuver and n-back types -



- Drivers are mo d decreases further as (2) t
- *None of the p ct of the length of a secondary

The length of secondary tasks affects driver interruptibility

1 2 **Getting darker (more likely to answer incorrectly for a given item)**



- Heat maps show the percentage of drivers who correctly answered each n-th item in the n-back task across maneuver and n-back types -

- Drivers are more likely answer incorrectly as (1) the item order (n-th) increases, and decreases further as (2) the demands of n-back type increase
- *None of the prior studies on driver interruptibility have either examined the impact of the length of a secondary task on interruptibility or considered the length in interruptibility

2.4. Discussion: Potential useful source for the classification of driver interruptibility

Speed pattern of drivers

Speed reduction patterns was varied according to maneuver types and driver interruptibility

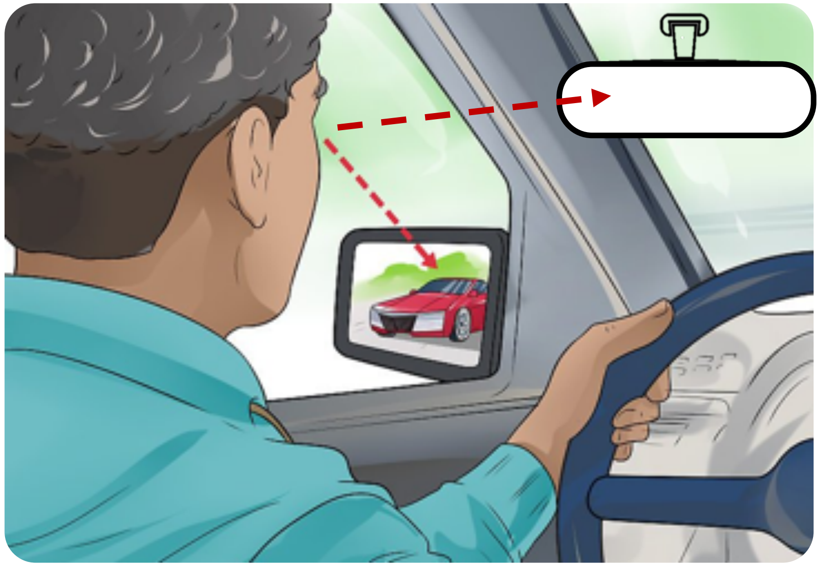
① Motion sensing and GPS tracking in smartphones [Bergasa 2014]

② Monitoring braking behaviors based smartphone cameras [Veeraraghavan 2007]

[1] Bergasa 2014. Drivesafe: An app for alerting inattentive drivers and scoring driving behaviors.
[2] Veeraraghavan 2007. Classifiers for driver activity monitoring.

2.4. Discussion: Potential useful source for the classification of driver interruptibility

Occurrence of subsidiary sub-tasks(e.g., checking mirrors)



[Researcher] *"When are the most difficult situation to engage in secondary tasks."*

[D14] *"Whether that car will be still there or move forward. And to estimate the speed of the car, I need to look in <u>the rear mirror</u>, as well as <u>the side mirror</u>, to hand the wheel, and to clarify the situation."*

*Prior studies considered major sub-tasks, such as braking, accelerating, and handling the steering wheel [Semmens 2019]

[1] Semmens 2019. Is now a good time?: An empirical study of vehicle-driver communication timing.

2.4. Discussion: Potential useful source for the classification of driver interruptibility

Occurrence of people (e.g., bicyclists or pedestrians) near the vehicle

[Researcher] *"When are the most difficult situation to engage in secondary tasks."*

[D15] *"When passing a crosswalk, <u>people</u> may cut in [...] I need to see what is coming, I get confused at answering, as I see the rear mirror, <u>bicycle</u> may come, see the side of the vehicle, and see the front of the vehicle."*

Study 3: Predicting driver interruptibility

Outlines

① **Key contributions**

② **Model building process**

③ **General model**

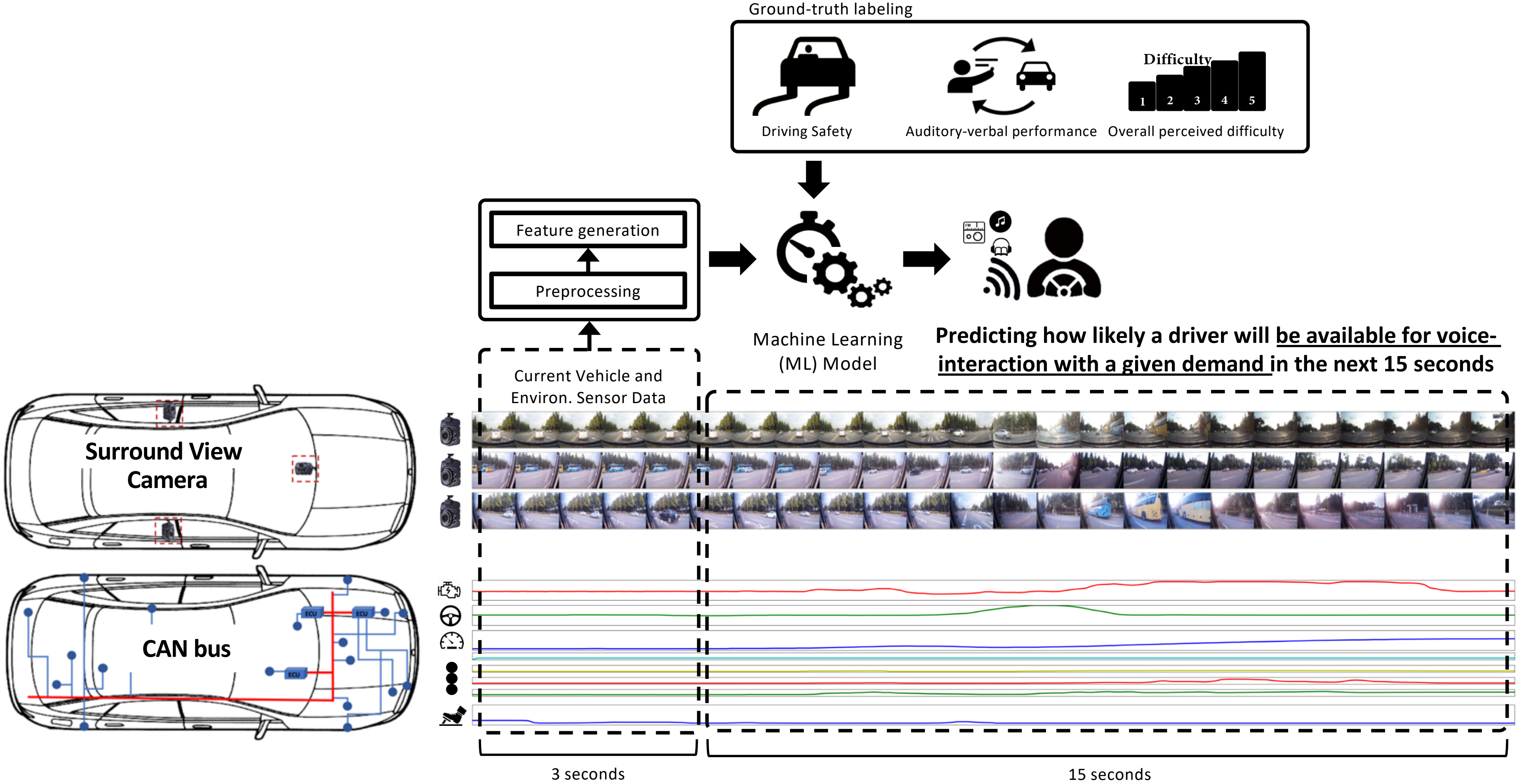
④ **Individual model**

⑤ **Discussion**

Kim A, Choi W, Park J, Kim K, Lee U. "Interrupting Drivers for Interactions: Predicting Opportune Moments for In-vehicle Proactive Auditory-verbal Tasks". In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. 2018 Dec 27;2(4):175.

3.1. Key contributions

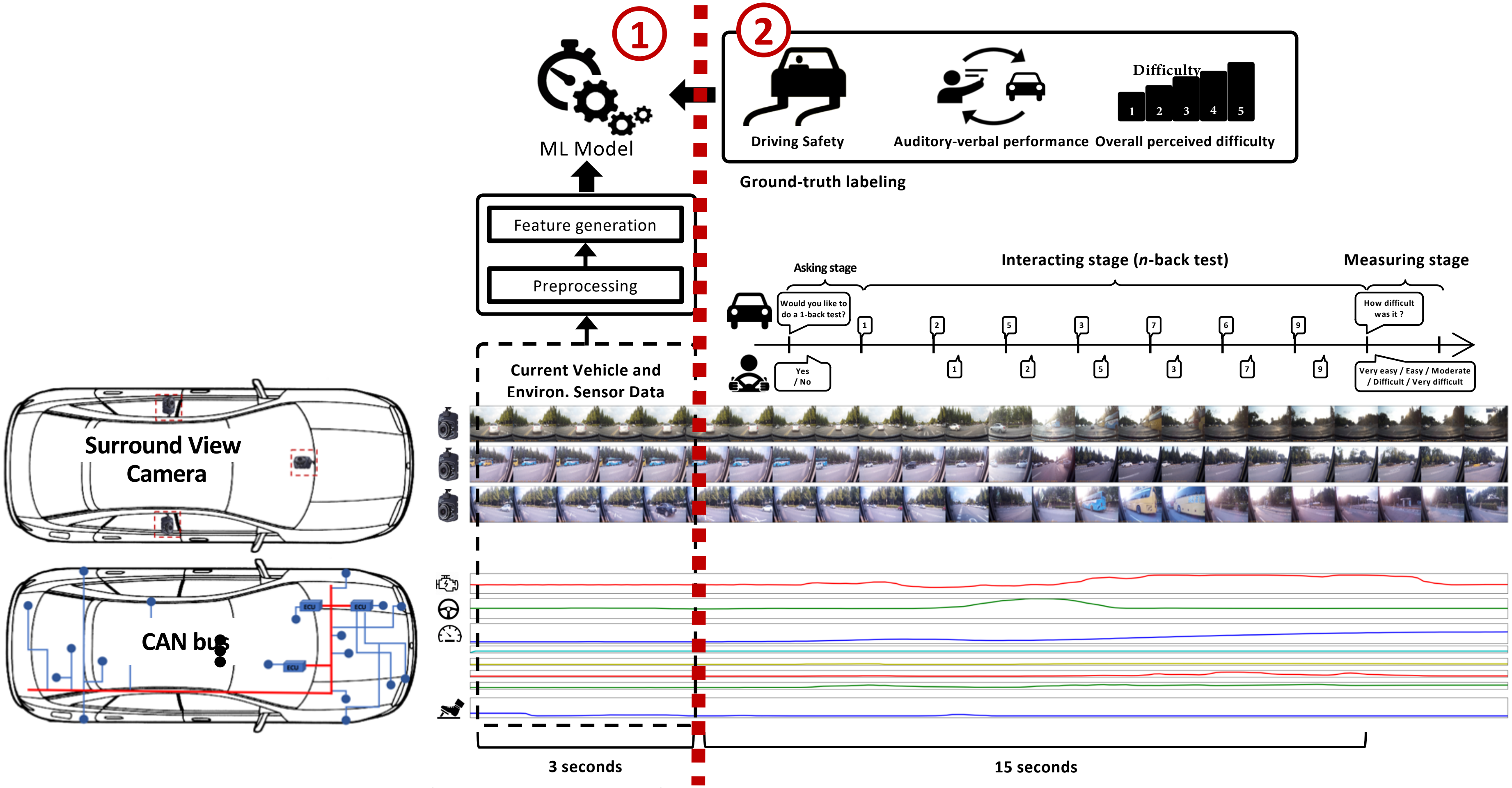
Developing models that predict driver interruptibility for a given demand of voice task (i.e., n-back types) based on the environmental and vehicular data prior to the voice task



Using diverse combinations of interruptibility definition and evaluated various machine learning models by considering varying window sizes, feature selection, and personalization.

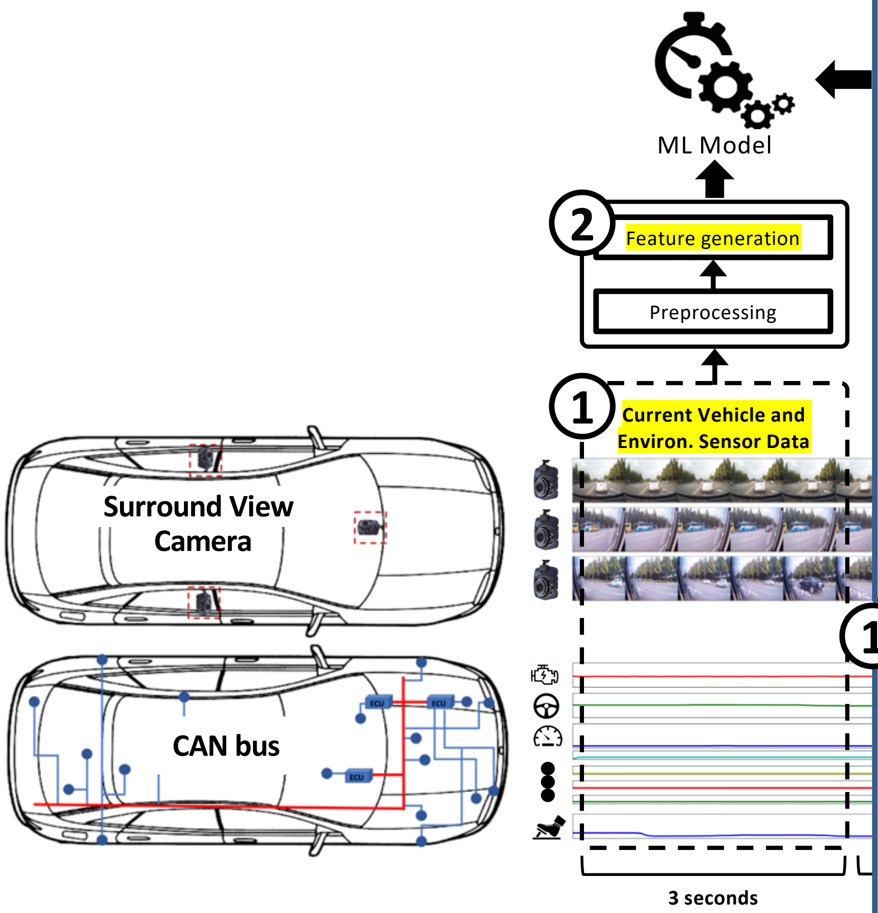
3.2. Model building process

Developing models that predict driver interruptibility for a given demand of voice task (i.e., n-back types) based on the environmental and vehicular data prior to the voice task

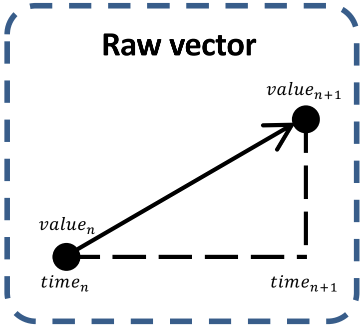


Vehicular and environmental data (3-second window) prior to the start of secondary task and the cognitive demand of an incoming secondary task (n-back type) were used for feature generation

3.2. Model building process: Feature generation



2



$$raw\_vector = \frac{value_{n+1} - value_n}{time_{n+1} - time_n}$$

$$abs\_vector = |vector|$$

$$pos\_vector = f(vector) \begin{cases} vector & \text{if } vector > 0 \\ none & \text{otherwisw.} \end{cases}$$

$$neg_{vector} = f(vector) \begin{cases} vector & \text{if } vector < 0 \\ none & \text{otherwise} \end{cases}$$

Calculate a set of features by taking two or more of the mathematical operators

Dataset	Mean	SD	Min	Max	Mdn	Skew.
Raw data	yes	yes	yes	yes	yes	yes
Raw vectors	yes	yes	yes	yes		
Absolute vectors	yes	yes				
Negative-valued vectors	yes	yes	yes	yes		
Positive-valued vectors	yes	yes	yes	yes		

1

Vehicle data

- Steering wheel angle
- Throttle position
- Engine position
- Accel. pedal position
- RPM
- Speed
- Turn indicator
- Brake pressure
- Accelerometer

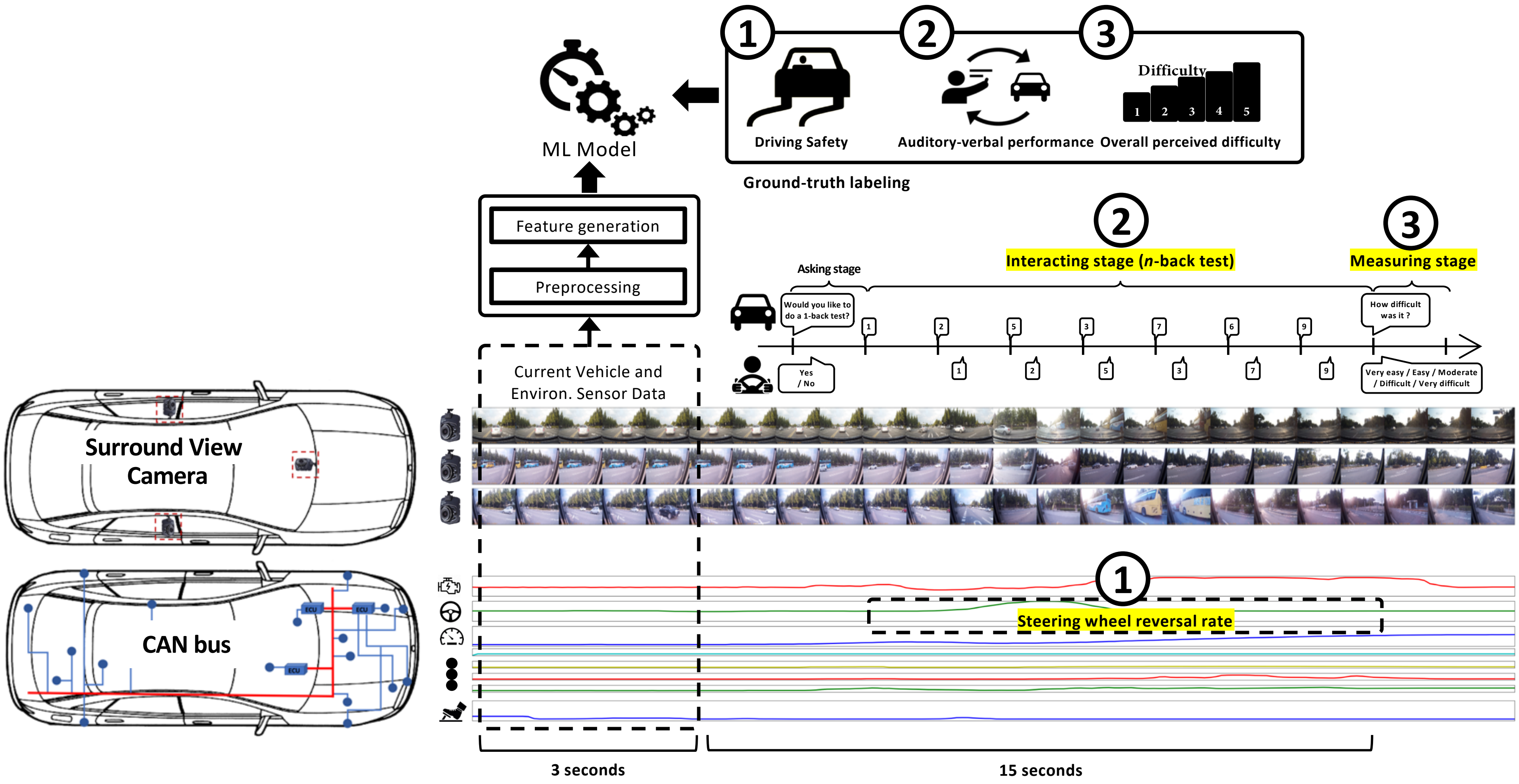
Driving environmental data

- Type of maneuver
- Number of vehicles in the front, left-lane, right-lane of the vehicle
- Distant to adjacent vehicles in the front, left-lane, right-lane of the vehicle

Voice interaction data

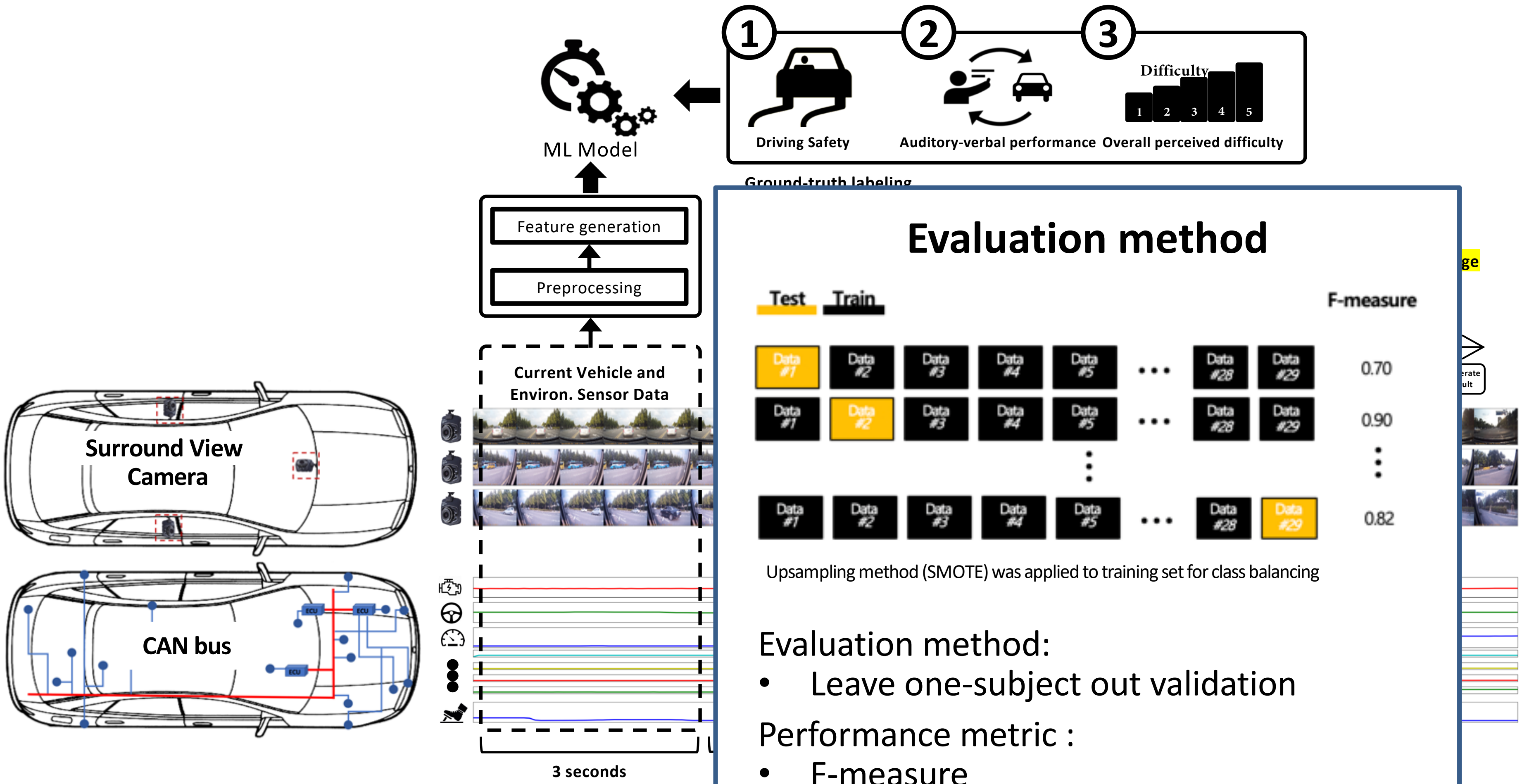
- *N*-back type

3.2. Model building process: Ground-truth labeling



3.2. Model building process: Evaluation method

Each model was trained with data of all drivers except one driver, and this driver's data was used to test the trained model



Upsampling method (SMOTE) was applied to training set for class balancing

Evaluation method:

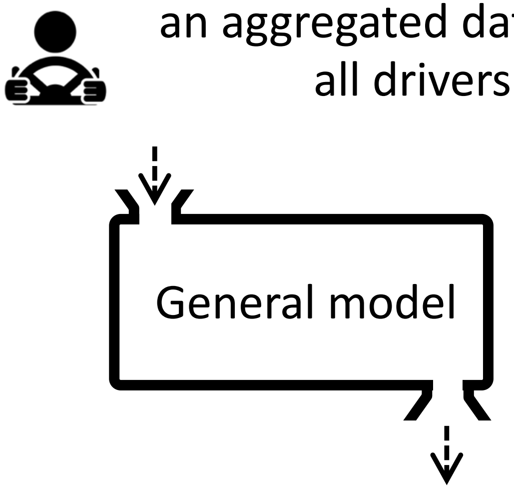
- Leave one-subject out validation

Performance metric :

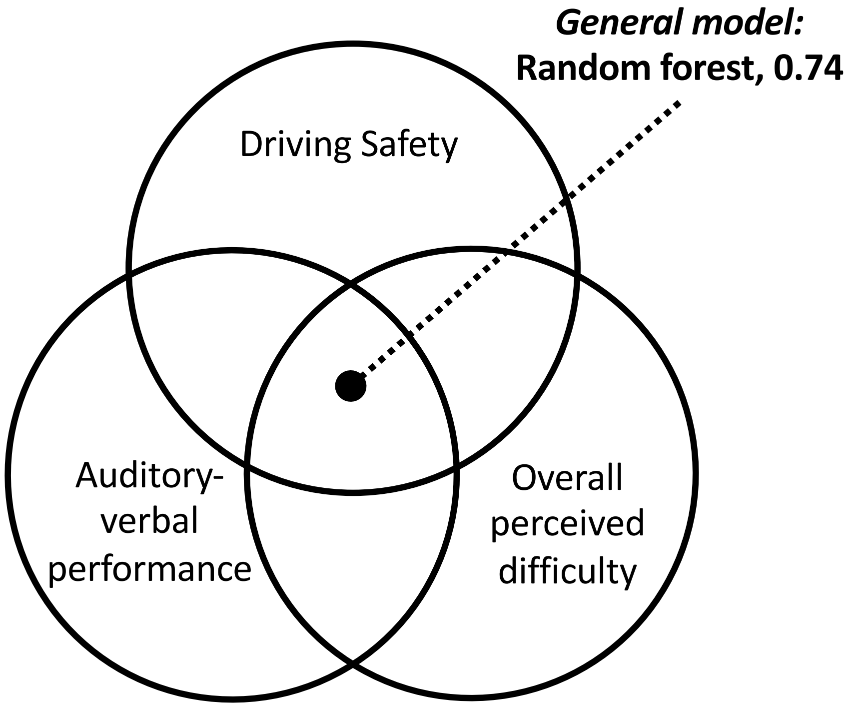
- F-measure

Building a model using an aggregated dataset of all drivers

an aggregated dataset of all drivers

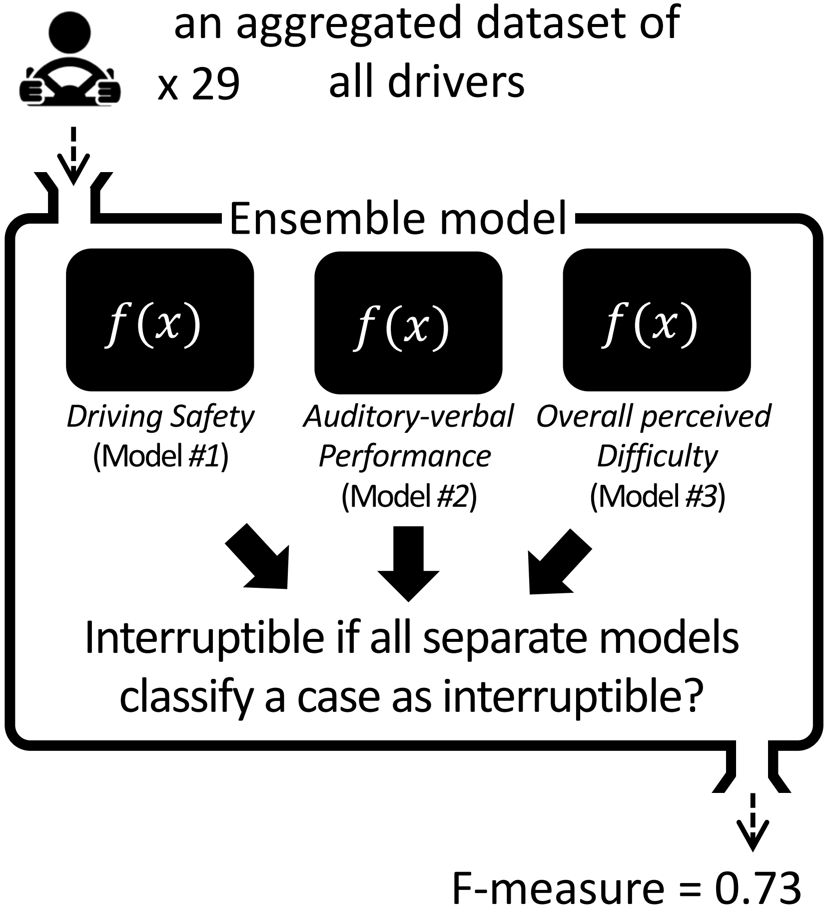


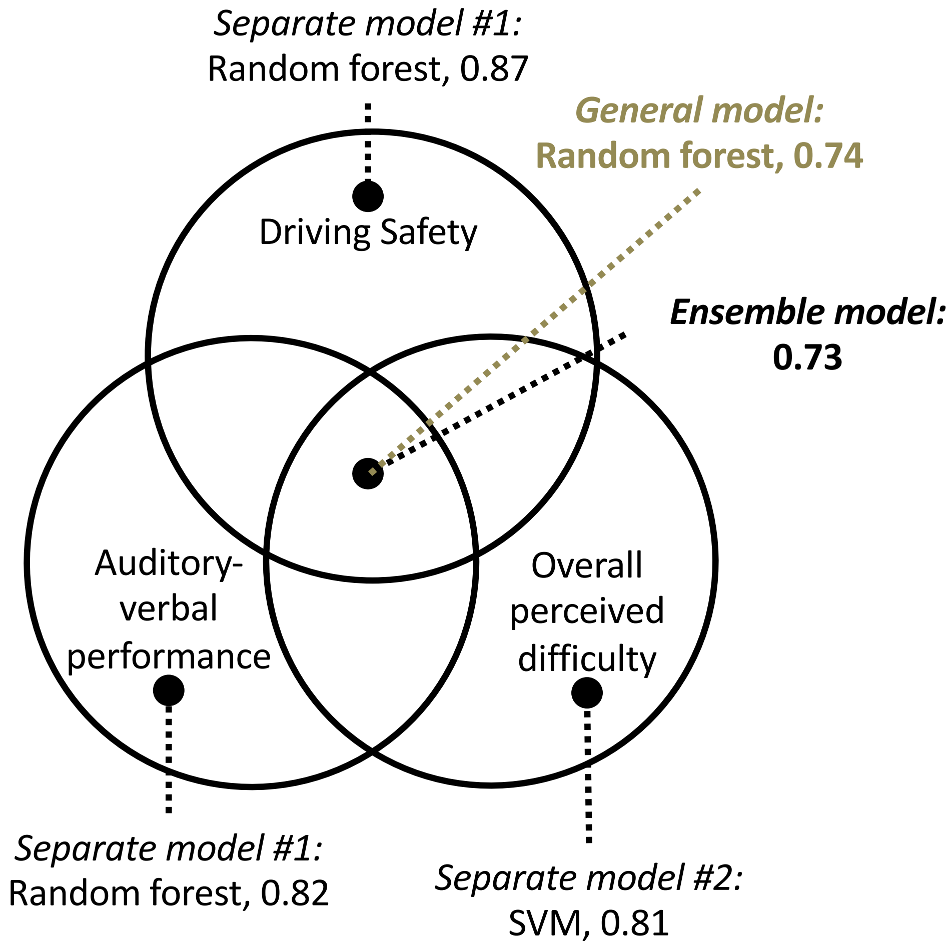
F-measure = 0.74



3.3. General model: Alternative Model Building Strategy (Ensemble model)

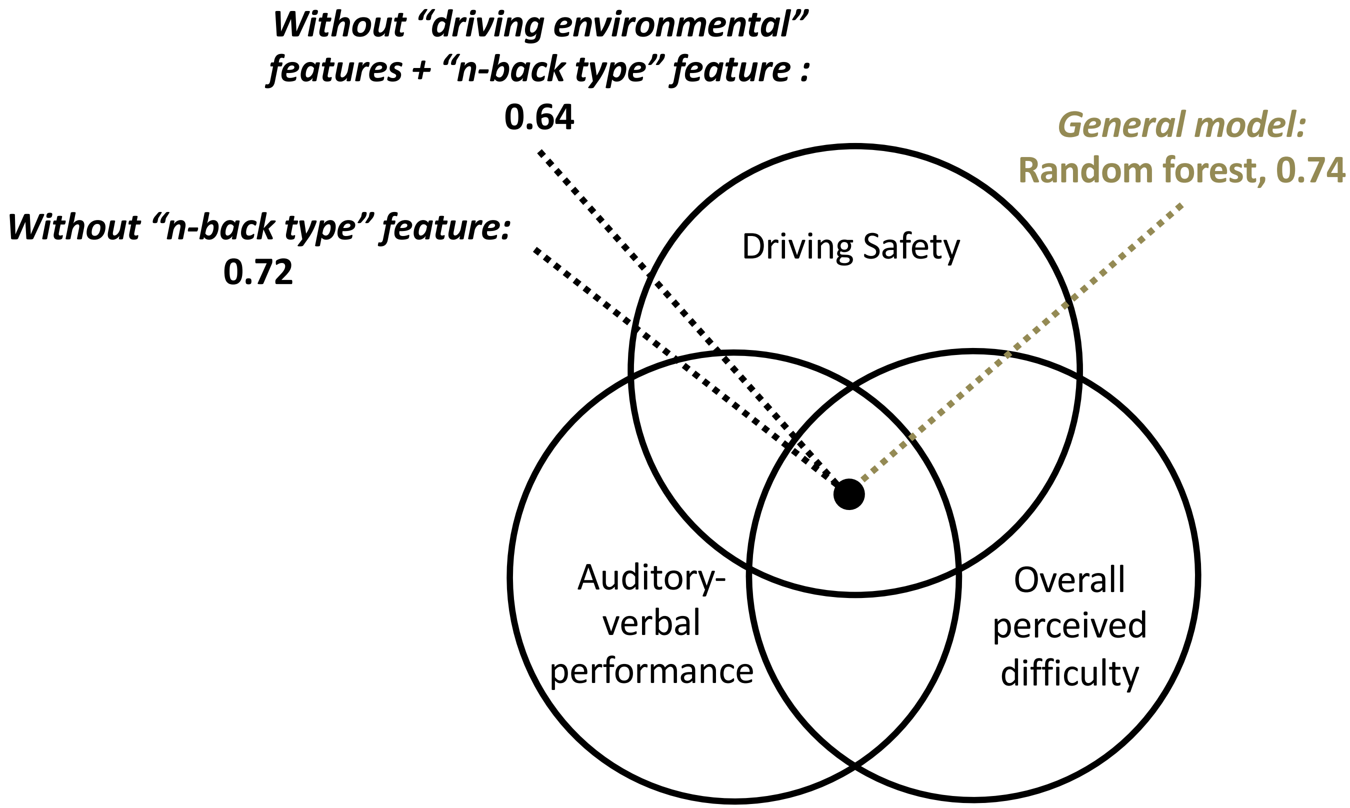
Three separated models that predict each of interruptibility dimensions, and the overall interruptibility can be predicted based on aggregating the results of the three models.





3.3. General model: Various data sources

Building models using an aggregated dataset of all drivers

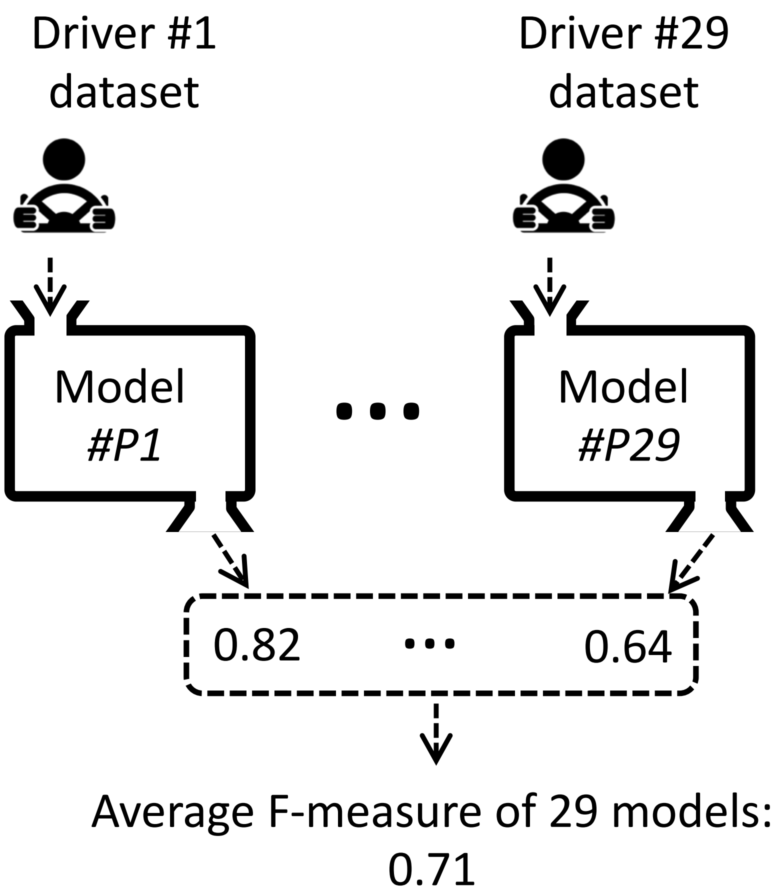


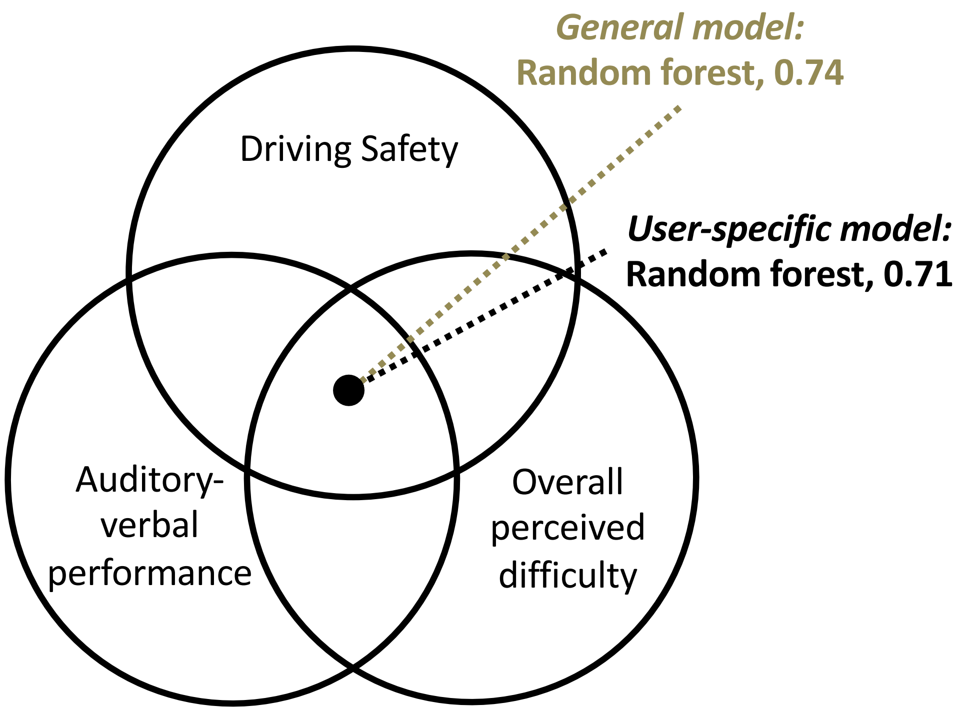
Driving environmental features

- Type of maneuver
- Number of vehicles in the front, left-lane, right-lane of the vehicle
- Distant to adjacent vehicles in the front, left-lane, right-lane of the vehicle

3.4. User-specific model

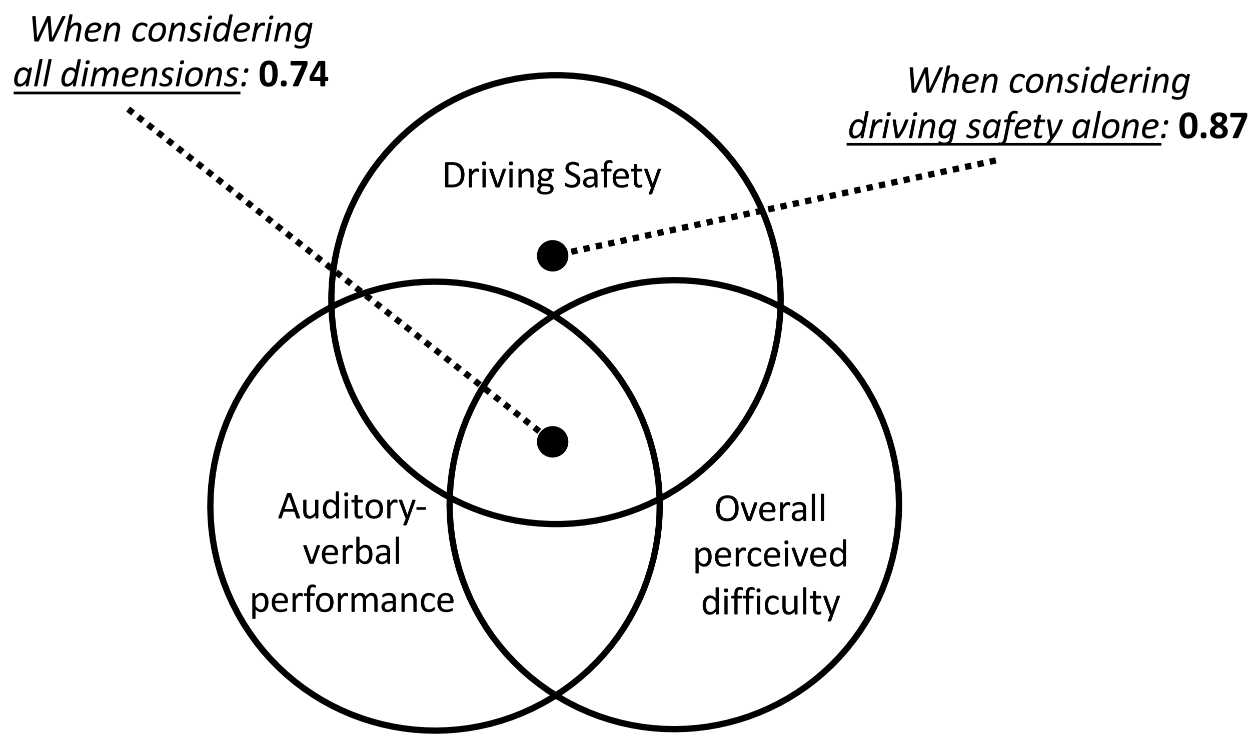
Individually trained and tested with specific user data





3.5. Discussion: Prediction of in-vehicle opportune moments

Interruptible moments can be reasonably inferred based on vehicular and environmental data

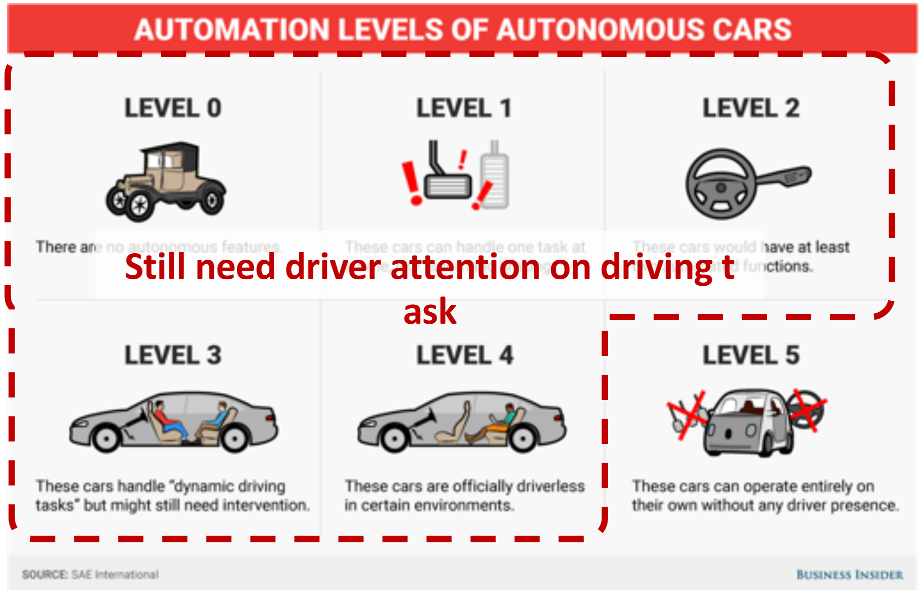


3.5. Discussion: Design implications

Facilitating In-vehicle Auditory-verbal Interactions

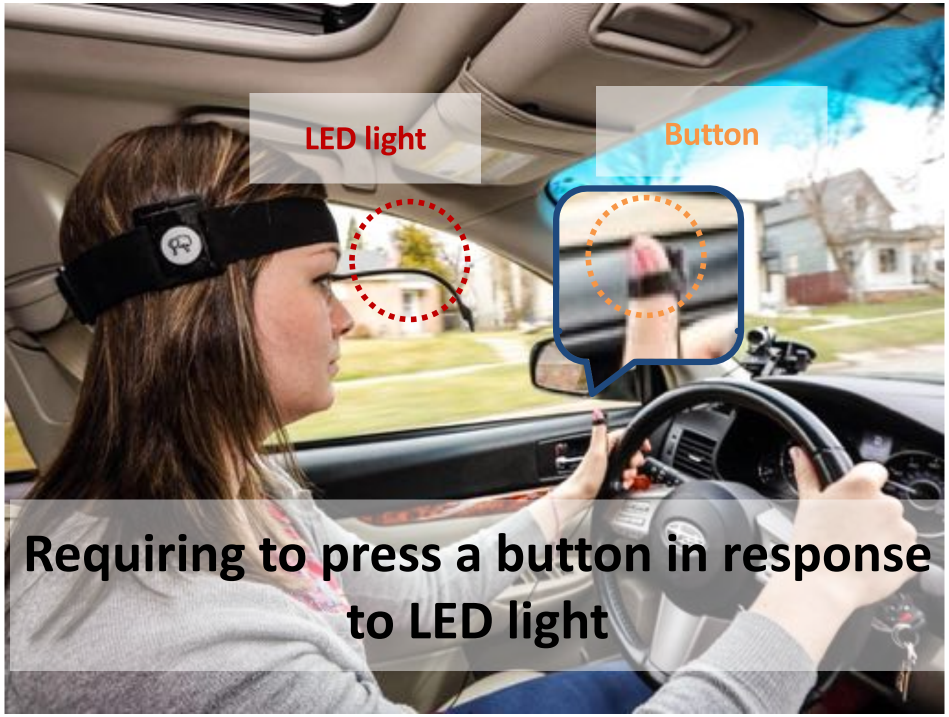
- To build a virtual agent that can proactively interact according to driver workload (e.g. initiating an interaction with drivers to overcome boredom and shake-off drowsiness in long-distance driving)
- To warn and proactively limit user-initiated interactions (e.g., verbally writing text) when drivers are classified as uninterruptible

3.5. Discussion: Driver interruptibility for self-driving car



- Until the level of self-driving cars reach LEVEL 5, <u>human drivers have to task part in the driving tasks from time to time</u> as, for instance, in a takeover or malfunction situations (e.g., vehicle not recognizing sudden appearance of pedestrians).
- In such situations, drivers may not be able to fully respond in a given situation if the drivers had been highly cognitively distracted (i.e., lack of **situation awareness)**.

3.5. Discussion: Driver interruptibility for self-driving car



- **Alternative metric for driving safety could be DRT – which measure driving-situation awareness.**
- **Measures of visual attention are strong indicators of distraction potential, reflecting that the crash risk increases with the amount of time a driver does not pat attention to the forward roadway.**
- **The DRT is a common task to measure visual attention of drivers**

[1] NHTSA 2011. Developing a Test to Measure Distraction Potential of In-Vehicle Information System Tasks in Production Vehicles.
[2] NHTSA 2014. Detection response task (DRT) evaluation for driver distraction measurement application.

Design Concept:
Towards Flow Control of Driver-Vehicle Voice Interactions

Outlines

① **A concept of flow control**

② **System-initiative flow control**

③ **User-initiative flow control**

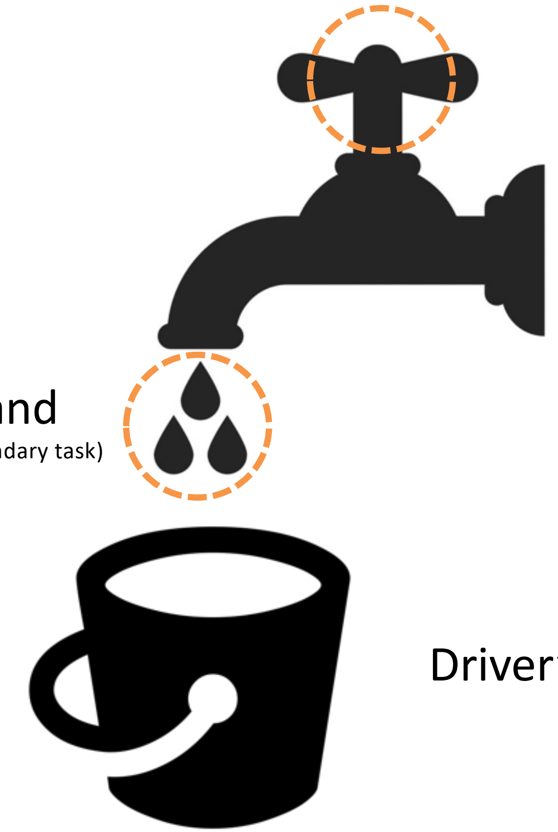
④ **Mixed-initiative flow control**

Kim A, Choi W, Park J, Kim K, Lee U. "Interrupting Drivers for Interactions: Predicting Opportune Moments for In-vehicle Proactive Auditory-verbal Tasks". In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies,* 2(4), 175. 2018 Dec 27

4.1. A concept of flow control

Managing voice interaction flows such that a system does not overwhelm a driver

Controlling flows according to interaction contexts



Driver-vehicle interaction contexts
(driving context and secondary-task demand)

Interaction demand
(Overall workload of driving and secondary task)

Driver’s limited capacity of cognitive resources
(Residual resources are depending on driving context)

4.2. Three types of flow control

Depending on WHO have control over the flows: System-, User-, or Mixed-initiative flow control

System-initiative flow control

- Proactive interactions are initiated and led by systems
- A system (system-initiative flow control) can fully automatically controls the flow of interactions by considering the interruptibility of their drivers.

Driver-initiative flow control (current in-vehicle interface)

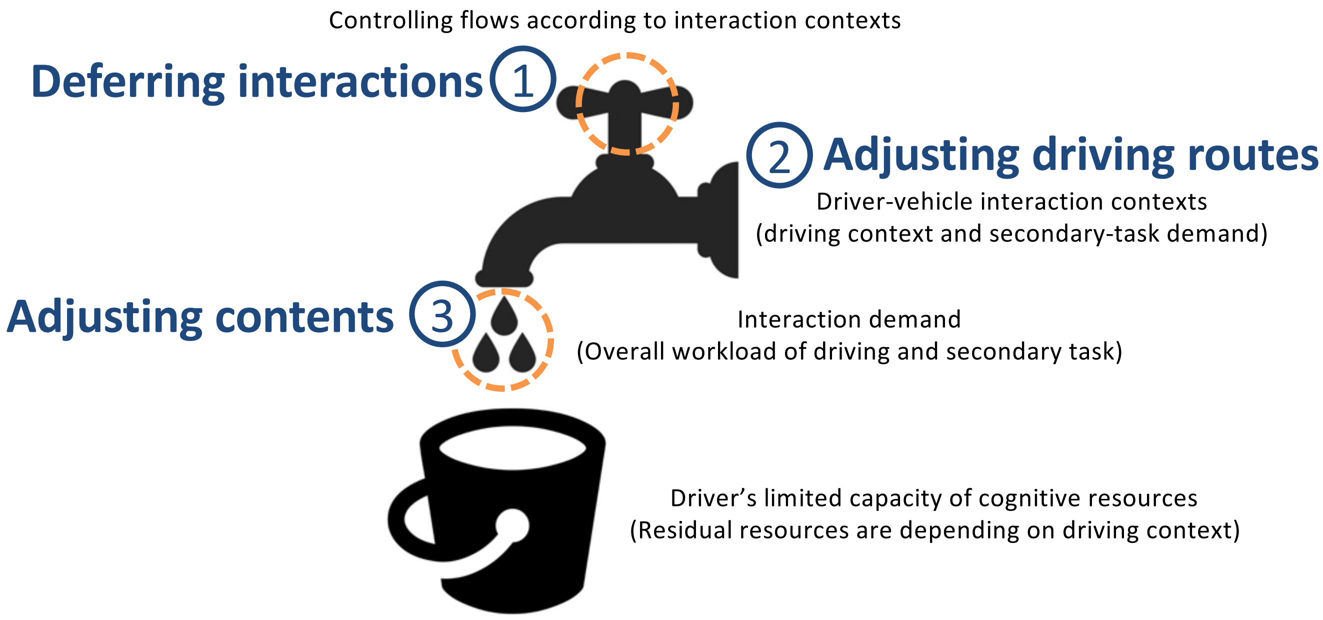
- Drivers have full control on the interaction flow (e.g., starting)
- But, lacking fine-grained, explicit flow controls like natural conversations (e.g., temporarily holding conversations until opportune moments appear).

Mixed-initiative flow control

- Driver and system collaboratively control the flow
- When there is high “uncertainty” in driver interruptibility, the system makes a joint-decision on whether to start interactions by asking drivers to resolve the uncertainty.

4.3. System-initiative flow control

Enforcing interactions to be initiated and terminated by tracking driver interruptibility



Systems can enforce and ensure interactions occur when a driver is interruptible in following approach:

① Deferring interactions until a driver become interruptible (e.g. interruptibility prediction [Study 3], speed pattern [study 2])

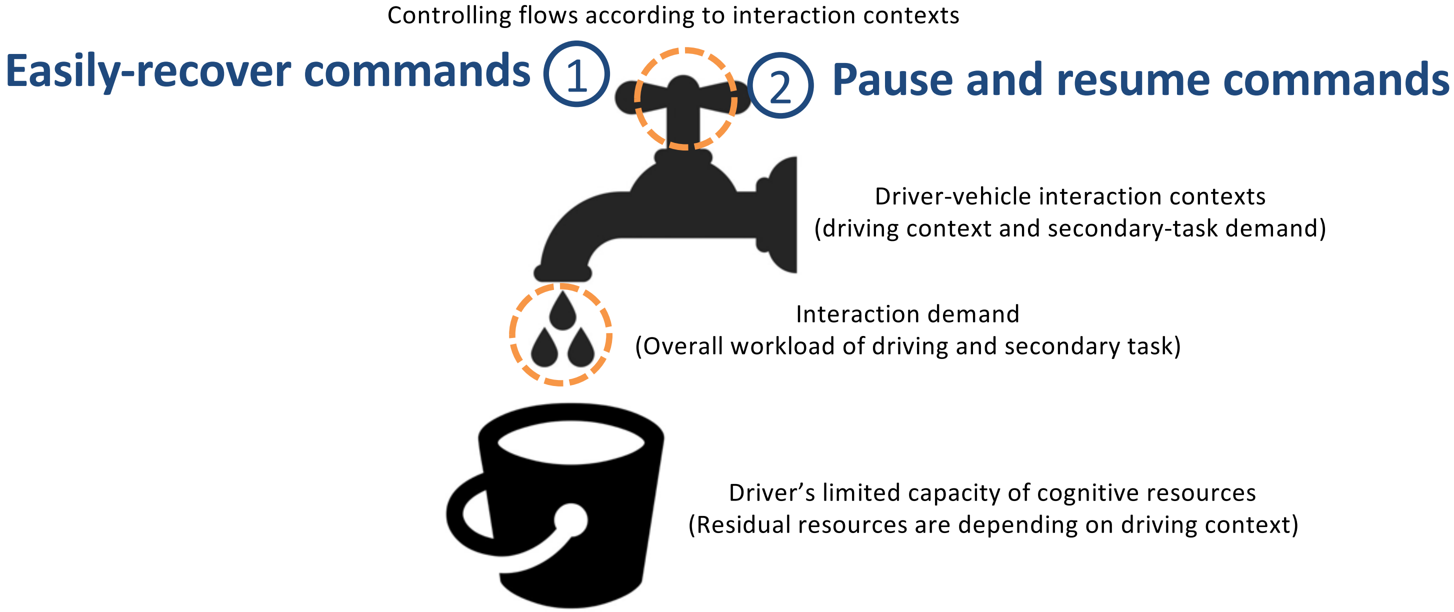
② Nudging the driver to adjust their driving workload (e.g. adjusting navigation routes [Study 2])

③ Dynamically adjusting the interaction workload (e.g. content summarization technique [Gambhir 2017], micro-interactions decomposition [Salton 1996])

[1] Salton 1996. Automatic text decomposition using text segments and text themes. HYPERTEXT '96
[2] Gambhir 2017. Recent automatic text summarization techniques: a survey. Artificial Intelligence Review 47

4.4. User-initiative flow control

Enhancing explicit means of controlling interaction flows



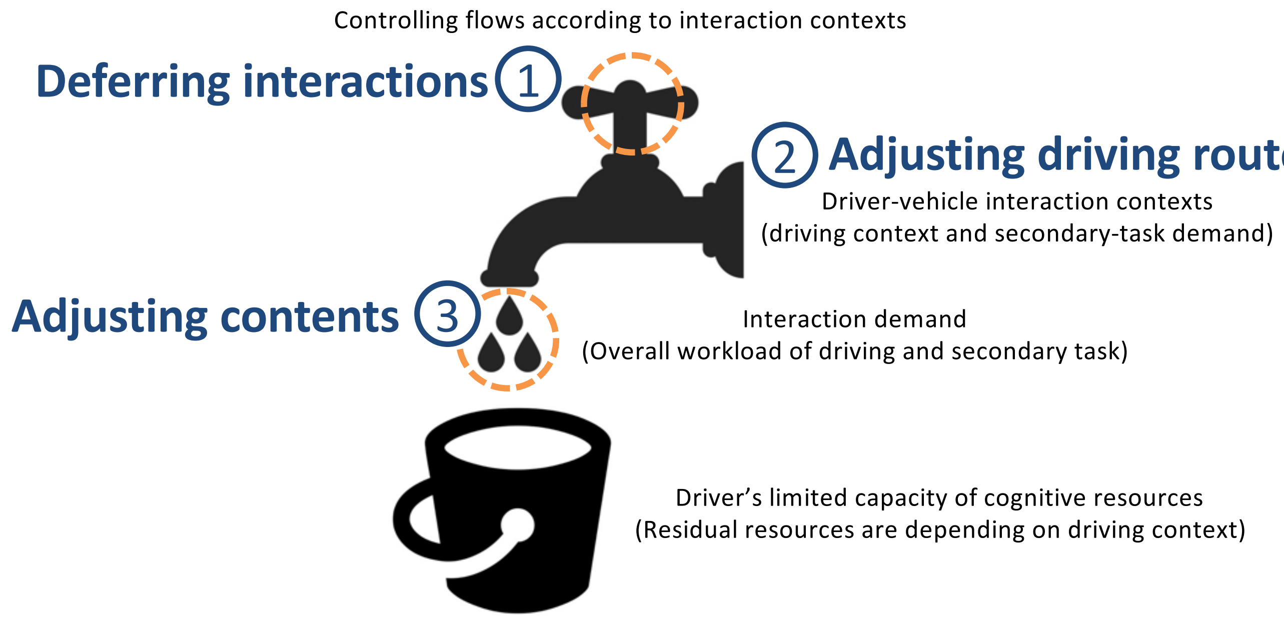
Drivers can control interactions using following naturalistic means of control driver-vehicle interactions:

① Enabling simple control commands, such as "Repeat" to easily recover after moments with a low auditory-verbal performance [Study 2].

② Let the drivers to explicitly control flow by enabling pause and resume commands such as "Hold on" and "Resume." [Study 2]

4.5. Mixed-initiative flow control

Making a joint-decision to revolve uncertainties (i.e. driver interruptibility)



Many uncertainties could arise in the classification of driver interruptibility (e.g., context dynamics, corner cases, and diversity of interruptibility measures). Systems can make a joint-decision to revolve uncertainties in following approach:

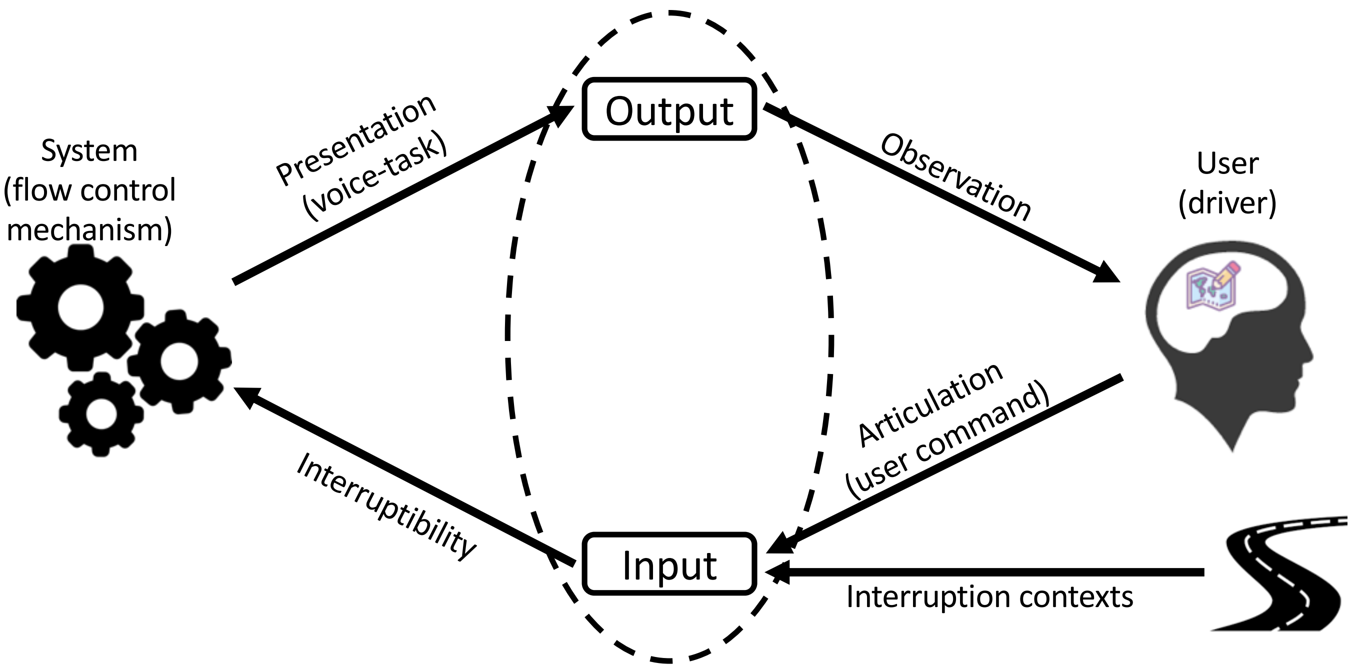
① Employing a dialog for user feedback, for example, “Would you like to engage in this service.”

② Hinting forthcoming interactions (e.g., by delivering advanced cueing/warning) based on driver behavior (Observing speed reduction pattern [Study 2]

③ Instead of auditory cure, employing a visual cue (e.g., showing the location of a forthcoming interaction in a navigation map) [Study 2]

4.6. Limitation and future study directions

Requiring follow-up studies on potential mismatch between user mental model and the actual system model



- For a given same user command, the output of flow control mechanism (e.g., timing and content of voice task) can be significantly varied by interruption contexts, leading a mismatch between the perceived mental model and the actual system model.
- Accumulated history of mismatches can cause an increase in complexity of use and/or a decrease in the frequency of use [Nothdurft 2016], and a decrease in the user's trust for the system [Muir 1994].

[1] Nothdurft 2016. Justification and Transparency Explanations in Dialogue Systems to Maintain Human-Computer Trust
[2] Muir 1994. Trust in automation: Part i. theoretical issues in the study of trust and human intervention in automated systems