Assessment of Hands-Free Outbound Texting and Navigation Destination Entry User Interface

Arnaud de Meulemeester
Managing Director
ATX International

ISO IEC ITU International Telecommunication Union

Research Objectives

- This research had two fundamental objectives
  - To assess task performance using an ATX developed speech-based interface
    - Texting tasks – relative to manual equivalent and baseline
    - Destination entry – relative to baseline
  - Assess learning required to use the speech-based interface
    - Examine and analyze naïve interaction
- Today's Presentation Focuses on Objective #1

Study Approach

- Learning Assessment
  - Provide a brief introduction to the system
    - What it can be used for
    - How to engage the system (Bluetooth)
  - Ask participant to complete 3 separate tasks
    - Texting & Destination entry
    - No further introduction or interaction by experimenter
  - Examine…
    - Success rate
    - Common errors
    - Sources of confusion
    - Time for users to be ‘comfortable’

Study Approach (continued)

- On-road Assessment
  - Participants perform 2 iterations of 10 tasks
    - Baseline (maintaining 45 mph)
    - Manual (Handheld) texting
      - Short, Medium, Long
    - Speech-based texting (identical to manual tasks)
      - Short, Medium, Long
    - Speech-based destination entry
  - Examine…
    - Task Outcomes (i.e., successful) and Durations
    - Vehicle measures (i.e., speed & lane maintenance)
    - Eye glance analysis
    - Workload Ratings (mental demand, frustration, awareness)

Participant Demographics

- 24 participants recruited from Southwest Virginia
  - New River Valley, Roanoke/Salem
- Two age groups, balanced by gender
  - 18-30 (younger); 45-55 (middle)
- Screening criteria
  - Must text weekly and own a smart phone; comfortable with touch screen keyboard
  - Have not participated in similar studies
  - Pass general health criteria

		N	Average	Std. Dev	Min	Max
Female	Younger	6	23.7	3.7	20	28
	Older	6	50.8	3.1	47	54
	Overall	12	37.3	14.6	20	54
Male	Younger	6	25.0	3.8	20	29
	Older	6	47.5	2.0	45	51
	Overall	12	36.3	12.1	20	51

Task List

- Baseline
- Handheld (manual) tasks
  - Texting
    - Short: “Testing 1 2 3” to Luke
    - Medium: “Have a nice day” to Dad
    - Long: “I’m driving to the grocery store” to Mary
- ATX (speech) tasks
  - Texting
    - Short: “Testing 1 2 3” to Luke
    - Medium: “Have a nice day” to Dad
    - Long: “I’m driving to the grocery store” to Mary
  - Destination
    - Address: 100 Clay Street, Blacksburg, VA
    - POI: Starbucks (closest location)
    - POI Category: Search for Gas Stations, choose Campus Exxon from the list

EYE GLANCE ANALYSIS

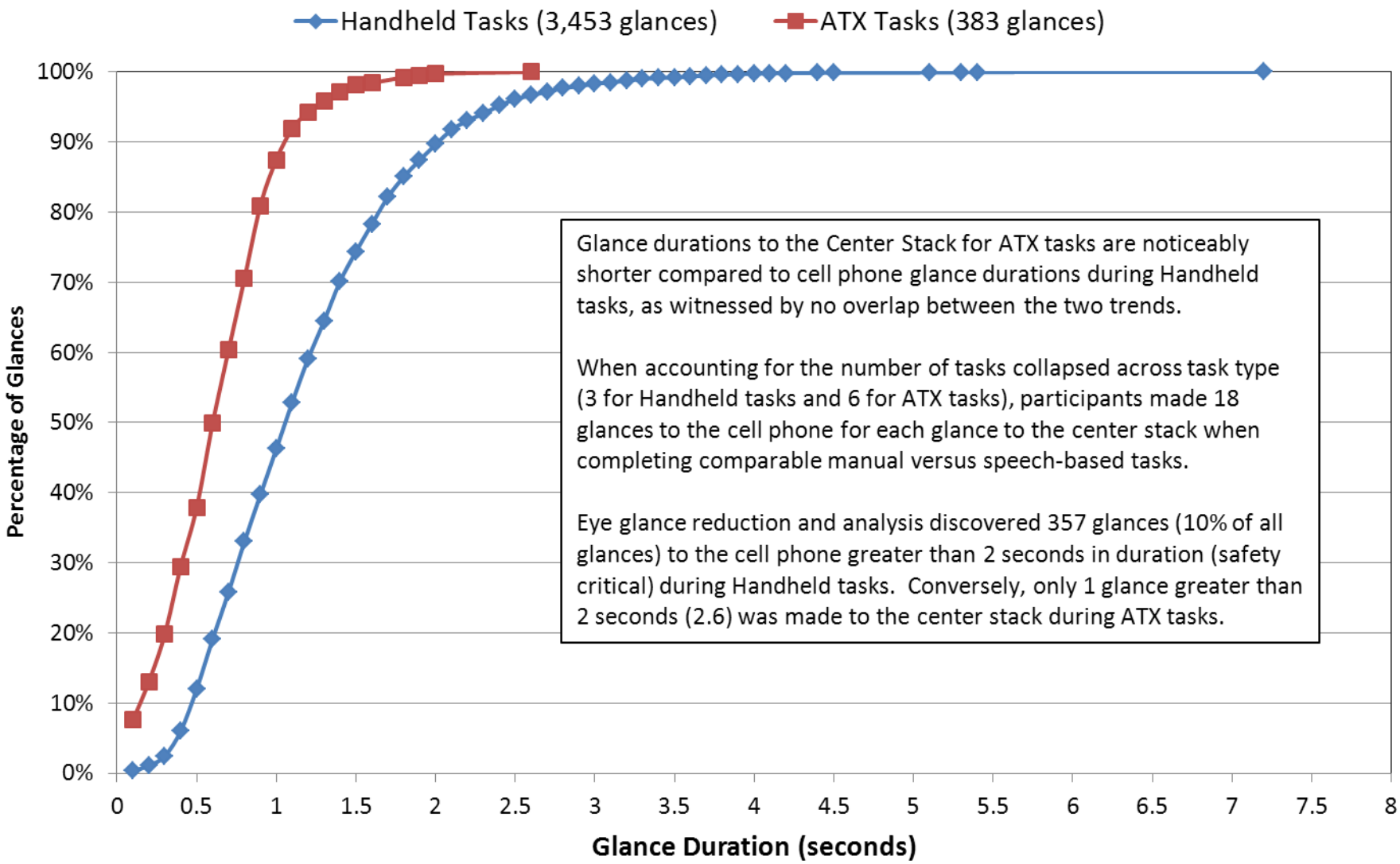
Eye Glance Analysis

- For task duration …
  - Frame by Frame eye glance reduction, including…
    - Forward
    - Instrument cluster (speedometer)
    - Mirror & shoulder checks (outside mirrors, rearview, OTS)
    - Center stack (Bluetooth module) – Source of Interest for speech-based tasks
    - Cell phone – Source of Interest for manual texting tasks

- Analysis
  - In the weeds…
    - Glance frequency by location
    - Glance duration by location
    - Sum of glance durations by location
  - What's important…
    - Percentage Eyes Off Road Time (%EORT)
      - %EORT = (Sum[all glances] – Sum[forward glances]) / (Sum[all glances])

Cumulative Frequency Distribution of Glance Durations to System of Interest Across Task Type

[System of Interest: Cell Phone for Handheld Tasks; Center Stack for ATX Tasks]

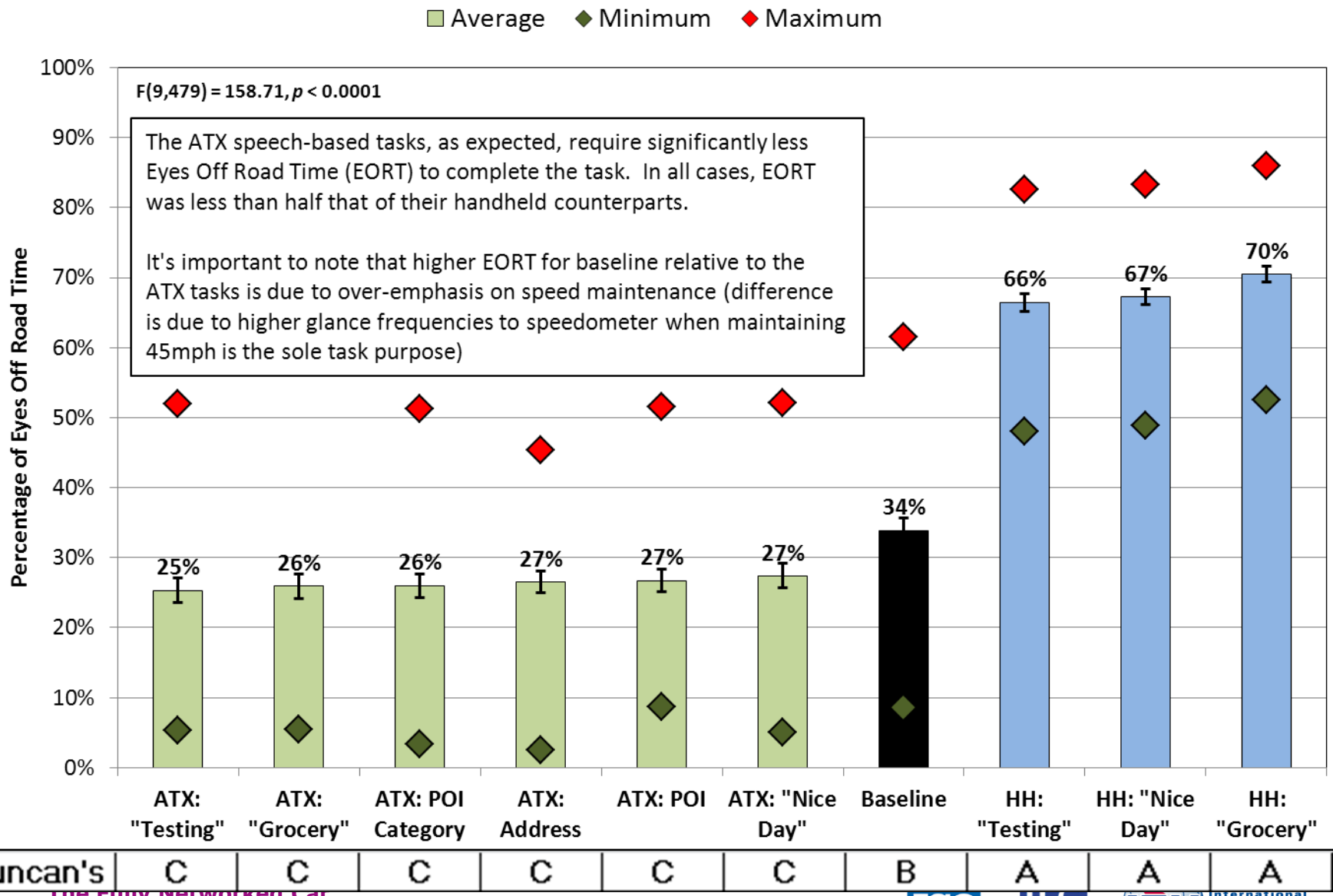


Glance durations to the Center Stack for ATX tasks are noticeably shorter compared to cell phone glance durations during Handheld tasks, as witnessed by no overlap between the two trends.

When accounting for the number of tasks collapsed across task type (3 for Handheld tasks and 6 for ATX tasks), participants made 18 glances to the cell phone for each glance to the center stack when completing comparable manual versus speech-based tasks.

Eye glance reduction and analysis discovered 357 glances (10% of all glances) to the cell phone greater than 2 seconds in duration (safety critical) during Handheld tasks. Conversely, only 1 glance greater than 2 seconds (2.6) was made to the center stack during ATX tasks.

Average Percentage of Eyes Off Road Time across All Participants and Trials

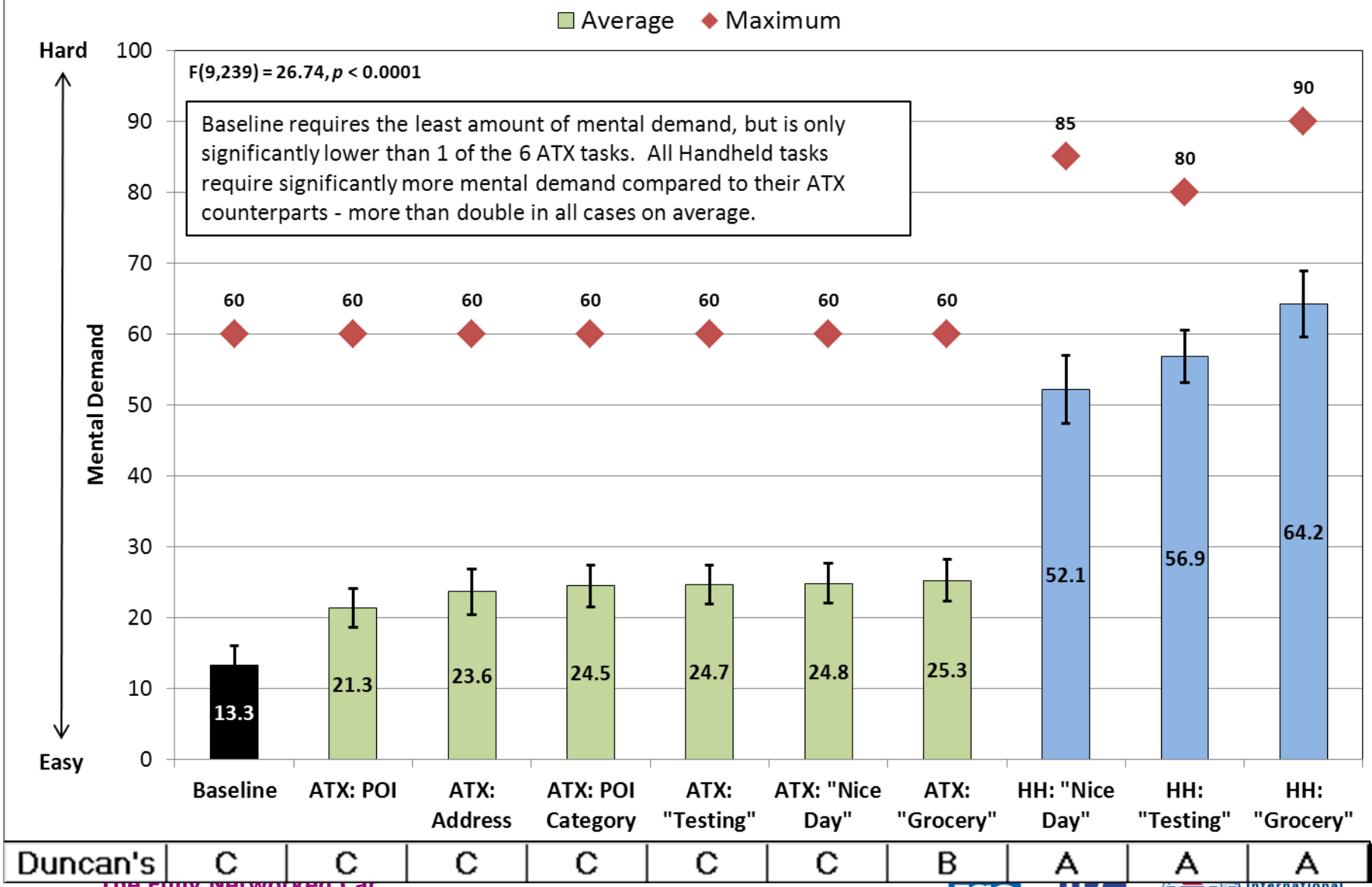


WORKLOAD RATINGS

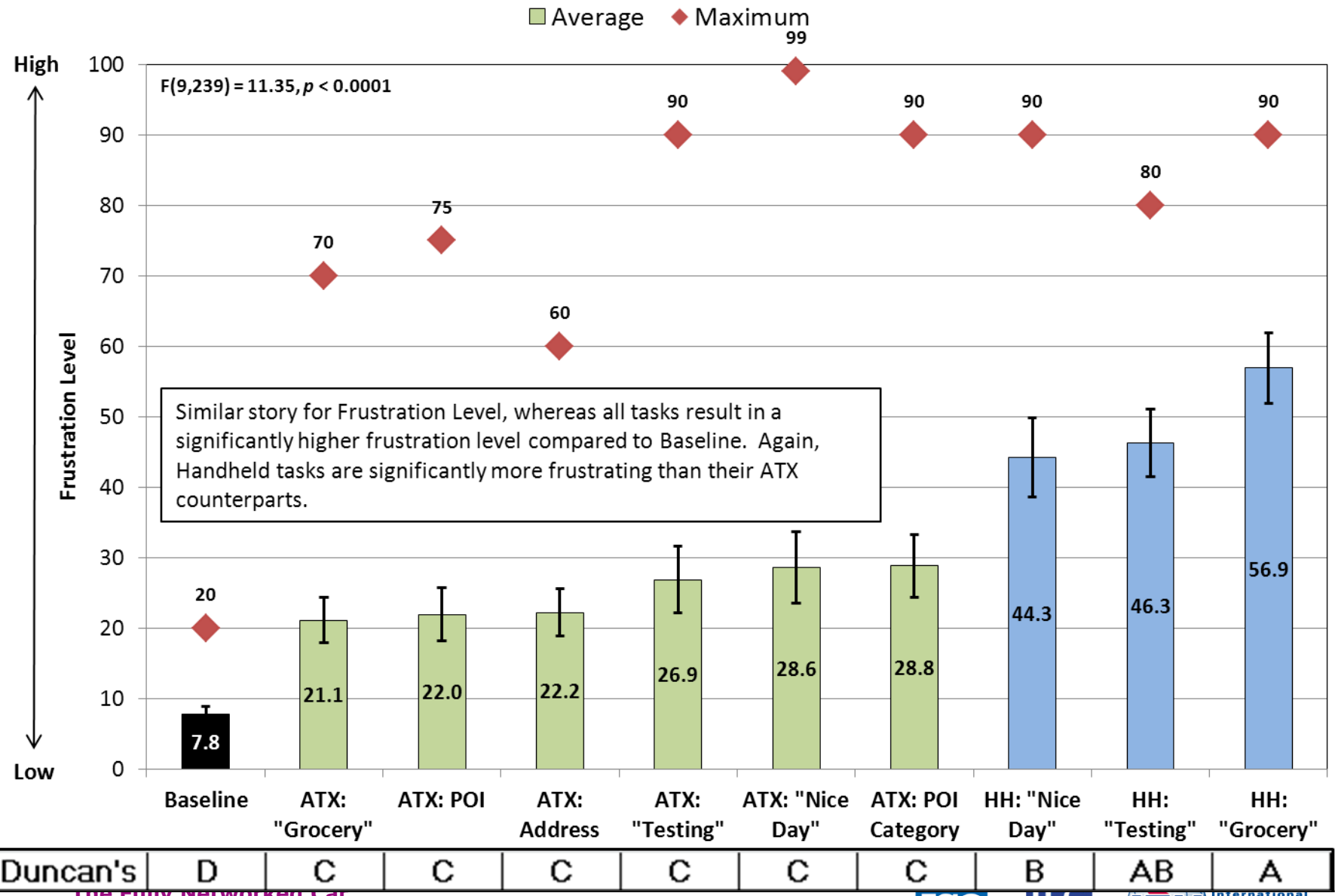
Workload Ratings

- Workload Ratings were asked following completion of 2nd trial of each task
  - Mental Demand – "…was task easy or demanding, simple, or complex?"
    - 1 (Easy) through 100 (Hard)
  - Frustration Level – "How … stressed, annoyed, versus … relaxed and complacent did you feel during the task?"
    - 1 (Low) through 100 (High)
  - Situation Awareness – "How aware were you of surrounding traffic when you were performing the task?..."
    - 1 (Low) through 100 (High)

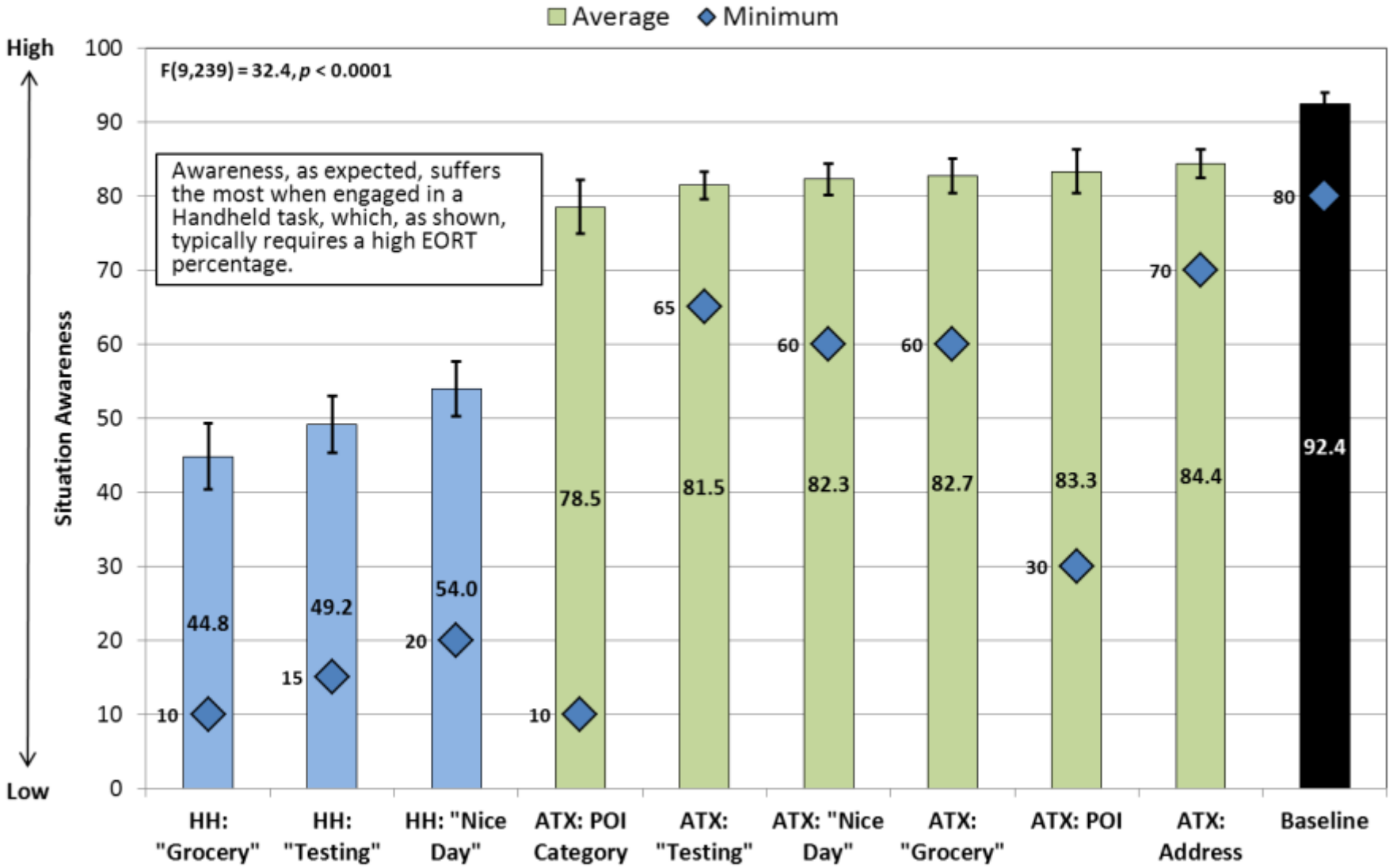
For driving safely and performing the task, what would be your rating for Mental Demand (N=24)



For driving safely and performing the task, what would be your rating for Frustration Level (N=24)



For driving safely and performing the task, what would be your rating for Situation Awareness (N=24)



Duncan's	D	CD	C	B	B	B	B	B	AB	A

True

False

POST DRIVE QUESTIONNAIRE

Post Drive Analysis

- Following On-Road assessment …
  - Participants were asked to complete a Post Drive Questionnaire

- Target Questions
  - Background info (prior experience with VR)
  - Ease of learning/use
  - Impact on safety
  - Appropriate conditions for use
  - Most beneficial; biggest concern; things to change
  - Sources of confusion
  - Desirability on personal vehicle

Post Drive Analysis - Overall

- Background…
  - 75% of participants noted some general experience with VR systems
    - 29% had in personal vehicle (i.e., OnStar, Bluetooth)

- Overall, participants…
  - found both systems intuitive and easy to learn
    - 96% for texting; 92% for destination entry
  - believe neither system distracts them from paying attention to the driving task
    - 88% for texting; 79% for destination entry
  - overwhelmingly liked both systems
    - 96% for texting; 92% for destination entry
  - felt both systems performed well
    - 92% for texting; 92% for destination entry

Post Drive Analysis – Open Ended

- Open Ended question summary on Texting interface (most frequently observed responses)…
  - What is the most beneficial feature(s) …?
    - 50% (12/24) answered that it allows you to keep eyes on the road while texting
    - 29% (7/12) also stated that it allows you to keep hands on the wheel while texting
    - 29% (7/24) felt it generally increased safety
  - What is your biggest concern…?
    - 58% (14/24) indicated accurate Voice Recognition
  - If you were designing this… how would you change it?
    - 38% (9/24) had no suggested changes
  - What do you feel is the most confusing aspect…?
    - 42% (10/24) stated that nothing about the interface was confusing
    - 25% (6/24) didn't always understand why system didn't work or recognize what they were saying

Post Drive Analysis – Open Ended

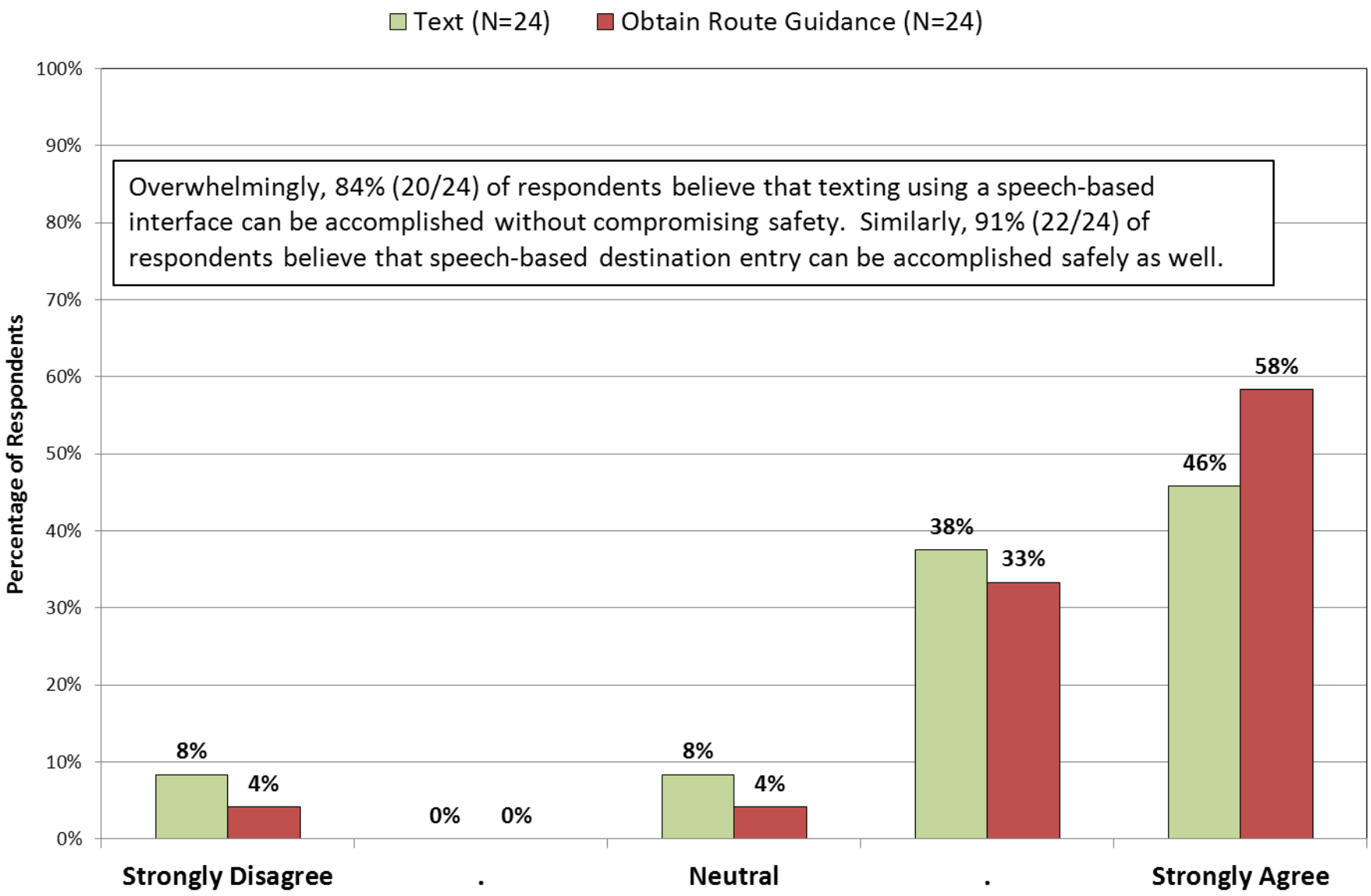
- Open Ended question summary on Destination interface (most frequently observed responses)…

  - What is the most beneficial feature …?
    - 25% (6/24) answered that feature allows user to keep hands on wheel
    - 25% (6/24) answered that feature allows user to keep eyes on road
    - 17% (4/24) indicated convenience the interface allows
  - What is your biggest concern…?
    - 29% (7/24) indicated no concerns
    - 25% (6/24) again were primarily concerned with VR accuracy
    - 21% (5/24) concerned about accuracy of database (receiving incorrect info)
  - If you were designing this… how would you change it?
    - 58% (14/24) had no suggested changes
    - 13% (3/24) thought users should be able to interject or skip (during lists)
  - What do you feel is the most confusing aspect…?
    - 54% (13/24) found nothing confusing

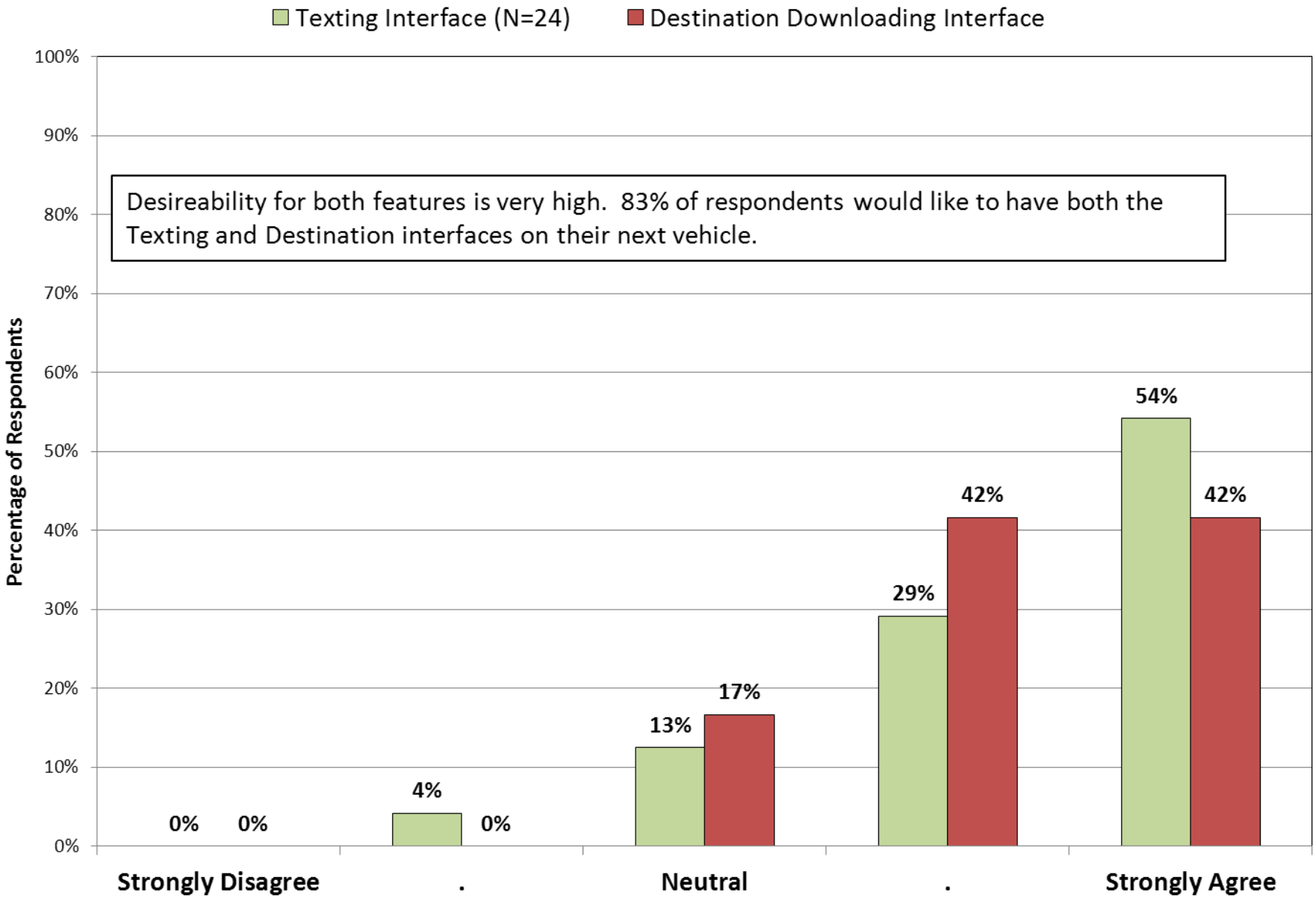
Post Drive Analysis – Open Ended

- Summary of infrequently made comments of interest (paraphrased)…
  - Texting Interface:
    - Hard to understand prompts
    - Would prefer to confirm contact before sending
    - If texting a lot may grow tired of hearing the formal prompts and revert back to manual texting
    - Could create a false illusion of total safety
    - Ask if message is correct before continuing
  - Destination entry:
    - Hard to understand prompts/options listed
    - Long list of options presented verbally (POI Category); Remembering list is ordered starting with nearest first
      - System should allow user to interject or skip when they know destination is not what they are looking for
    - Allow adding city to POI search criteria

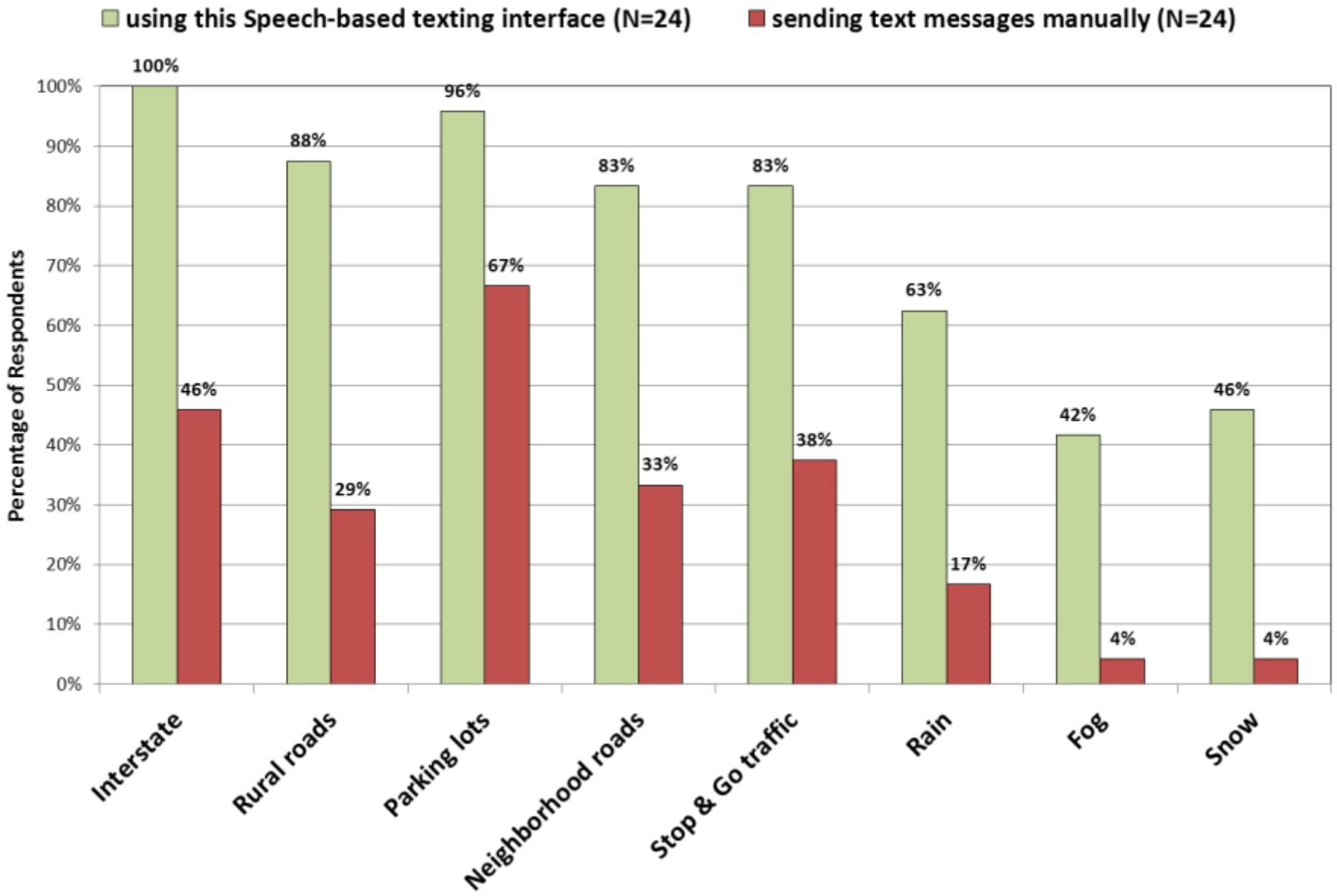
This speech-based interface allows me to ..... without compromising safety while driving.



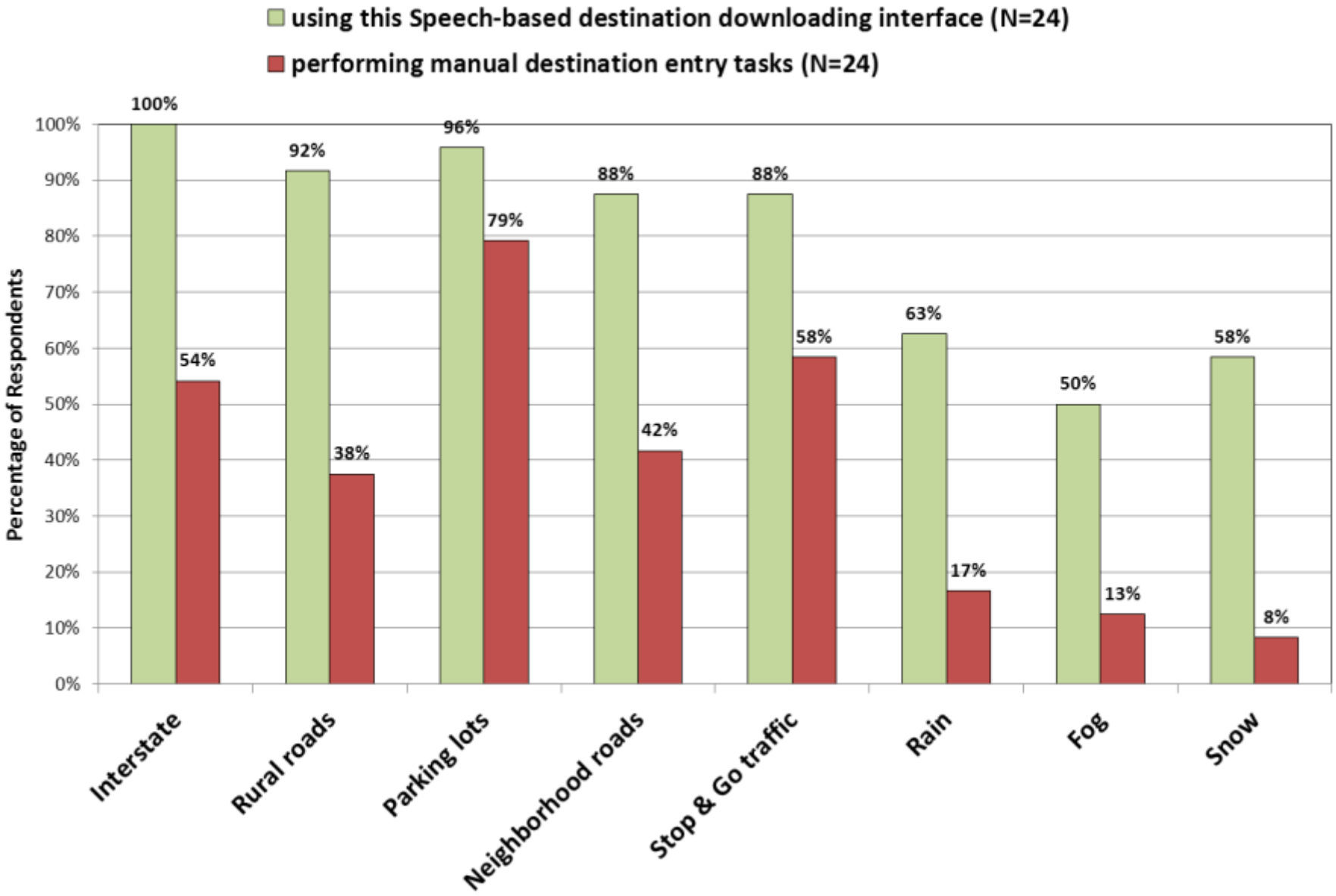
I would want to have this speech-based ..... feature on my next vehicle.

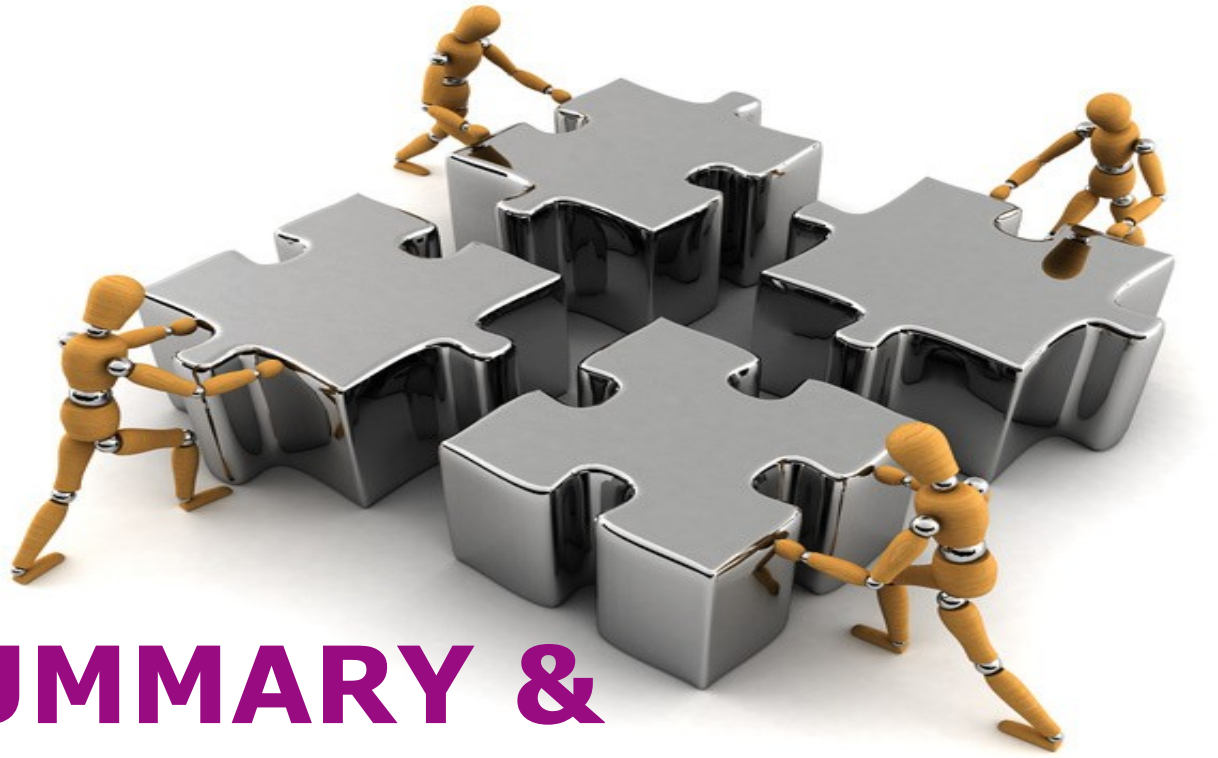


Under what environments and conditions would you feel comfortable...



Under what environments and conditions would you feel comfortable...





SUMMARY & CONCLUSIONS

Summary

- Clear results found related to both objectives
- Learning Assessment (Objective #2)
  - High overall task success rate during 1st attempts with minimal confusion
    - If confused, typically cleared up by end of task
- On-Road Assessment (Objective #1)
  - High success rates across all tasks; Manual (Handheld) vs. Speech-based (ATX)
  - No clear differences across vehicle network measures (speed, steering, etc.)
  - Lane deviations and eye glance analysis show significant advantages of speech-based tasks over manual [pending completed lane deviation reduction]
  - Subjective workload ratings favored ATX
- Post-Drive Questionnaire
  - Favorable; high desirability expressed across most questions

Conclusions

- Both objectives were achieved
  - Learning Assessment identified system ease of use by naïve users with no instruction
  - On-Road Assessment quantified anticipated advantages of speech-based tasks over their manual counterparts
    - Subjectively, high desirability of this interface, given a lack of training and exposure, coupled with most participants experiencing at least some voice recognition issues, is very positive

Arnaud de Meulemeester

Managing Director
ATX International
ameulemeester@atxg.com

Dr. Tom Schalk

Vice President of Voice Technology
ATX Group
tschalk@atxg.com