A Preliminary Assessment of Perceived and Objectively Scaled Workload of a Voice-Based Driver Interface

Bryan Reimer, Ph.D.
MIT AgeLab & New England University Transportation Center

Driver Assessment
Bolton Landing, NY

June 20, 2013

Correspondence > Bryan Reimer, Ph.D. > (617) 452 – 2177 > reimer@mit.edu

Acknowledgments

- Co-authors
  - Bruce Mehler
  - Hale McAnulty
  - Daniel Munger
  - Alea Mehler
  - Enrique Abdon Garcia Perez
  - Thomas Manhardt
  - Joseph Coughlin
- Sponsors
  - Santos Family Foundation
  - Region One New England University Transportation Center
  - Toyota's Collaborative Safety Research Center

THE SANTOS FAMILY FOUNDATION

MIT University Transportation Center

CSRC Collaborative Safety Research Center TOYOTA

Correspondence > Bryan Reimer, Ph.D. > (617) 452 – 2177 > reimer@mit.edu

Literature on the Demands of Voice Interfaces

While **A LOT** of research has been conducted on experimental and "wizard of oz" voice interactions, as well as handheld systems, more limited research address the demand of embedded production level vehicle systems:

- Carter and Graham (2000) – Jaguar S-Type (n=32; tracking task)
- Chang et al. (2005) – 2004 Accord & 2005 Acura RL (n=10x2; field)
- Harbluk et al. (2007) – 2005 Acura-TL (n=16 x 2; LCT)
- Shutko et al. (2009) – 2008 model year SYNC (n-25; sim)
- Owens et al. (2010a, b)* - 2010 Mercury Mariner with SYNC (n=21; field / test track)

*Only research with use of physical production integrated system

Correspondence > Bryan Reimer, Ph.D. > (617) 452 – 2177 > reimer@mit.edu

Research Questions

While voice interfaces offer the promise of reducing the time a driver's eyes are drawn away from the roadway, a number of questions remain:

- How do we effectively assess the amount of non-visual demand associated with voice interfaces?
- In what conditions does voice control reduce demand over traditional methods of interaction?
- How do different simple vs. more complex voice interactions affect drivers?
- Do age and gender impact drivers' perceptions and use of voice interfaces?

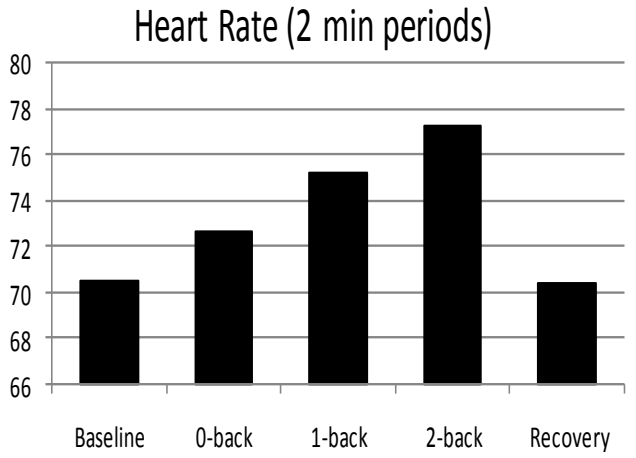
Correspondence > Bryan Reimer, Ph.D. > (617) 452 – 2177 > reimer@mit.edu

The MIT n-back

An Emerging International Method for Inducing Graded Cognitive Workload

- Series of 10 single digit numbers (0-9) presented in random order aurally at 2.25 sec intervals
- Subject instructed to respond with nth digit back
- Across levels
  - Auditory demands constant
  - Vocal demands "relatively" constant
- Aims to manipulate secondary cognitive demand

Stimulus	**6 9 1 7 0 8 4**
0-back Response	6 9 1 7 0 8 4
1-back Response	- 6 9 1 7 0 8
2-back Response	- - 6 9 1 7 0

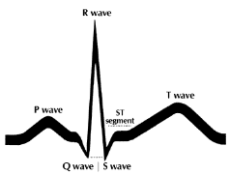


(Mehler, Reimer Dusek & Coughlin, 2011)

Correspondence > Bryan Reimer, Ph.D. > (617) 452 – 2177 > reimer@mit.edu

Instrumented Vehicle

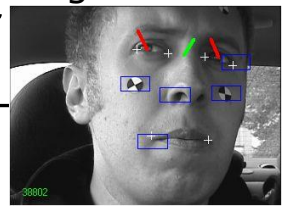
Physiology (HR, GSR)



Lane tracking

Audio recording

Cameras for monitoring driver & traffic



Eye tracking

GPS

PC based data acquisition system

Rear seat for overseeing experiment

Vehicle telemetry via CAN bus link

Correspondence > Bryan Reimer, Ph.D. > (617) 452 – 2177 > reimer@mit.edu

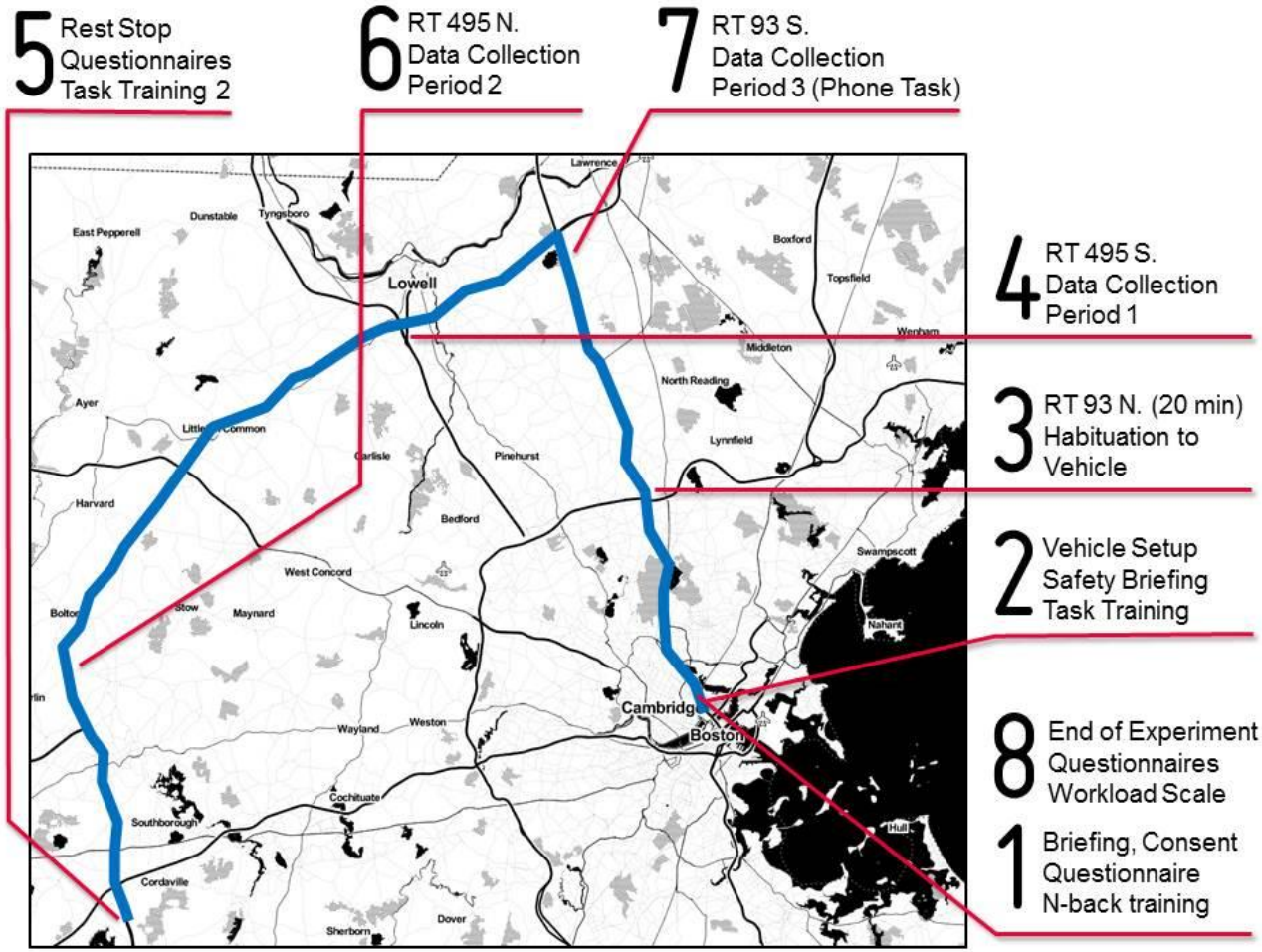
Interface Tasks

Extensive parking lot training and driving evaluation (x2)

- Visual manual task (radio tuning)
  - Single press preset selection
  - CAMP style manual radio tuning (Angell et al., 2006)
- Voice interface tasks
  - Preset selection (manual preset selection equivalent)
  - Tuning to a station (manual radio tuning equivalent)
  - Full address destination entry
  - Cancel navigation
  - Song selection (success and failure (1 replication))
  - Contact dialing

Correspondence > Bryan Reimer, Ph.D. > (617) 452 – 2177 > reimer@mit.edu

Experimental Route and Key Protocol Steps



Correspondence > Bryan Reimer, Ph.D. > (617) 452 – 2177 > reimer@mit.edu

Subjects

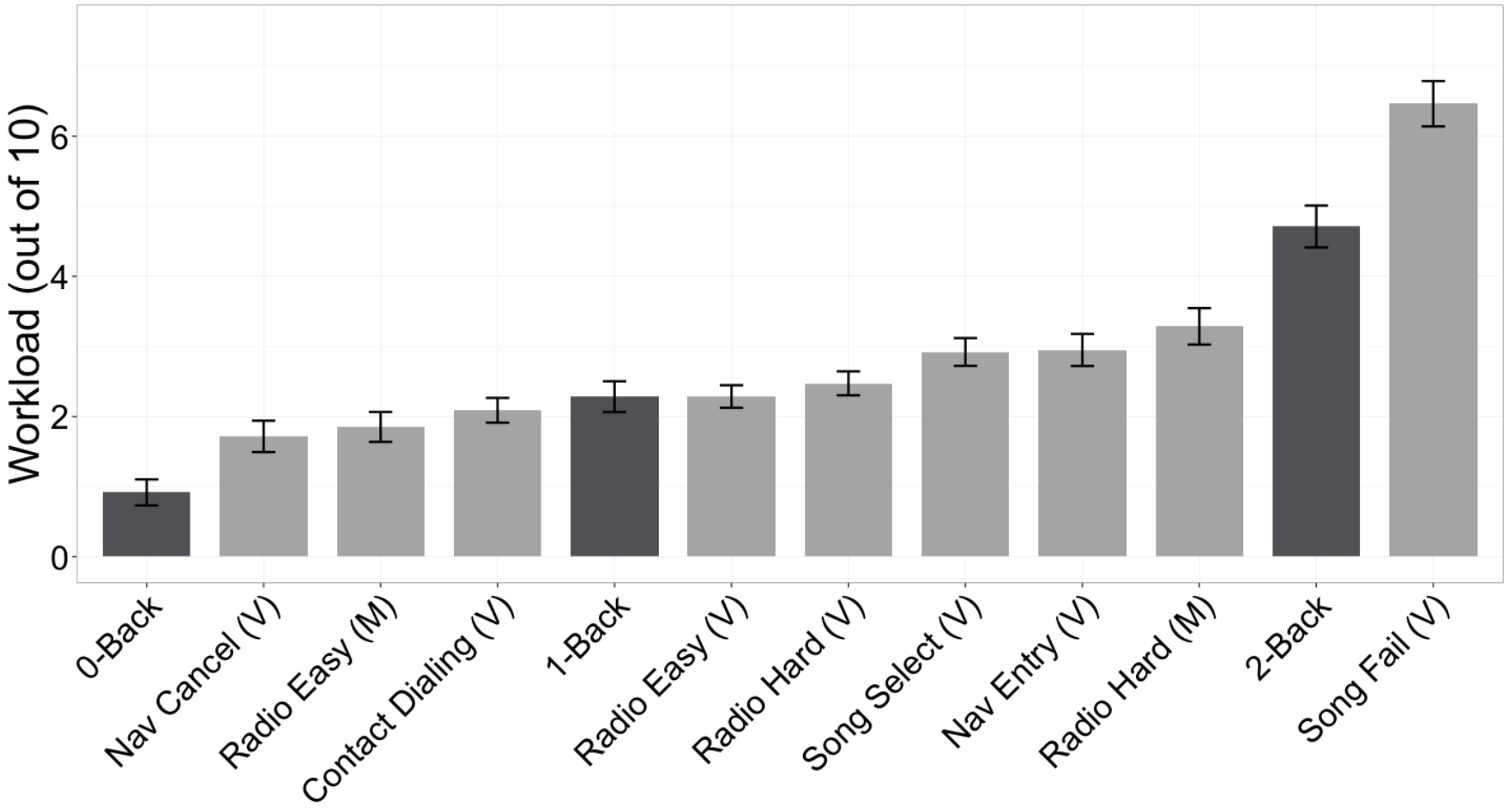
Two age groups: 20-29 and 60-69 years; 102 recruited

N	Reason for Exclusion
14	Unreliable/unusable ECG (heart rate) recording
6	Subject did not pass MoCA screening
6	Equipment failure
4	Scheduling problems prevented completion of study
4	Subject demonstrated unsafe driving behavior
4	Subject was part of the pilot version of the study
2	System did not recognize subject's voice (determined in parking lot before drive)
1	Unsafe conditions on road
1	Research assistant executed experiment protocol incorrectly
42	**Total Excluded**

60 case analysis sample equally distributed across age and gender

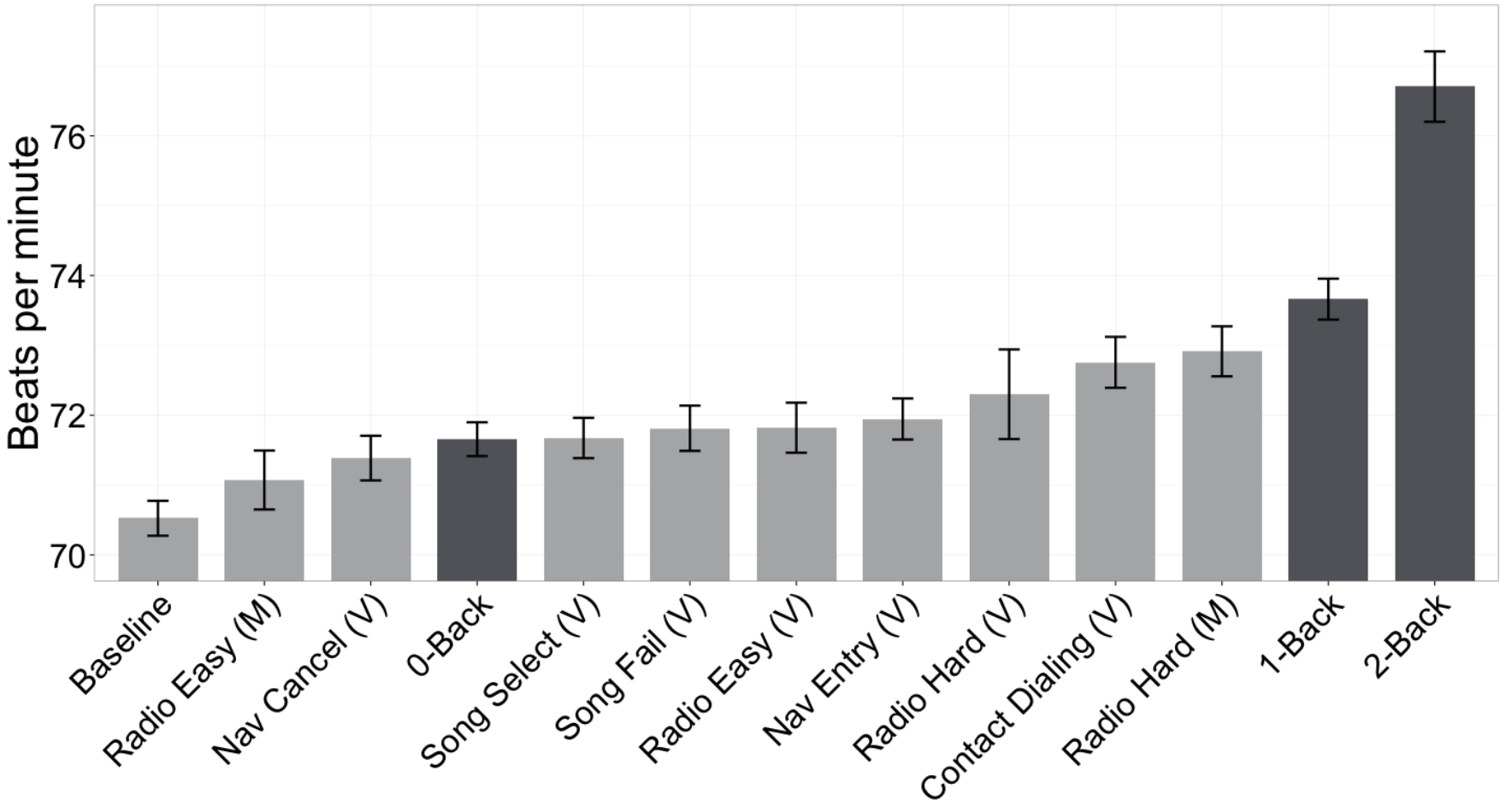
Correspondence > Bryan Reimer, Ph.D. > (617) 452 – 2177 > reimer@mit.edu

Self-Reported Workload



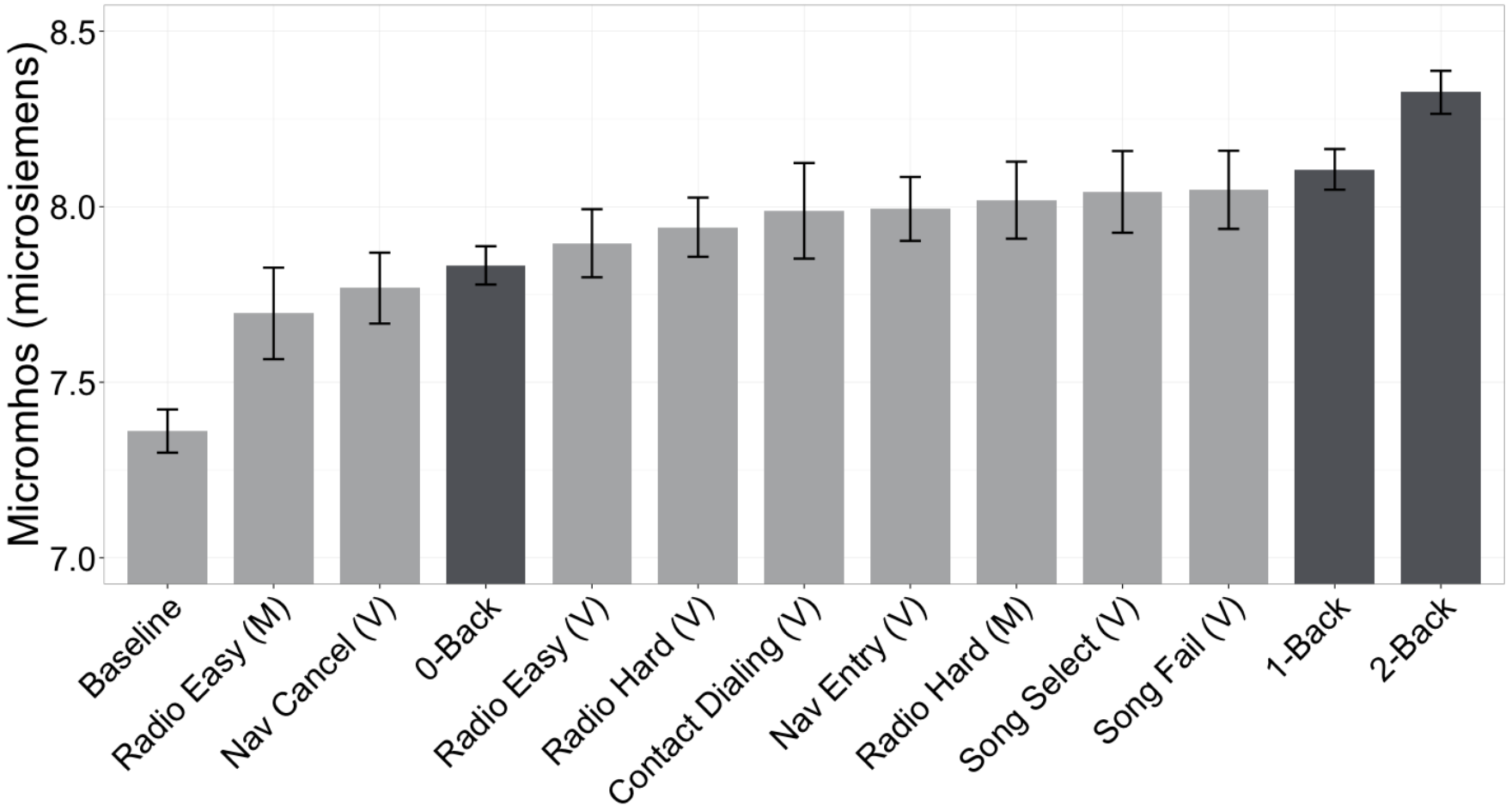
Correspondence > Bryan Reimer, Ph.D. > (617) 452 – 2177 > reimer@mit.edu

Heart Rate



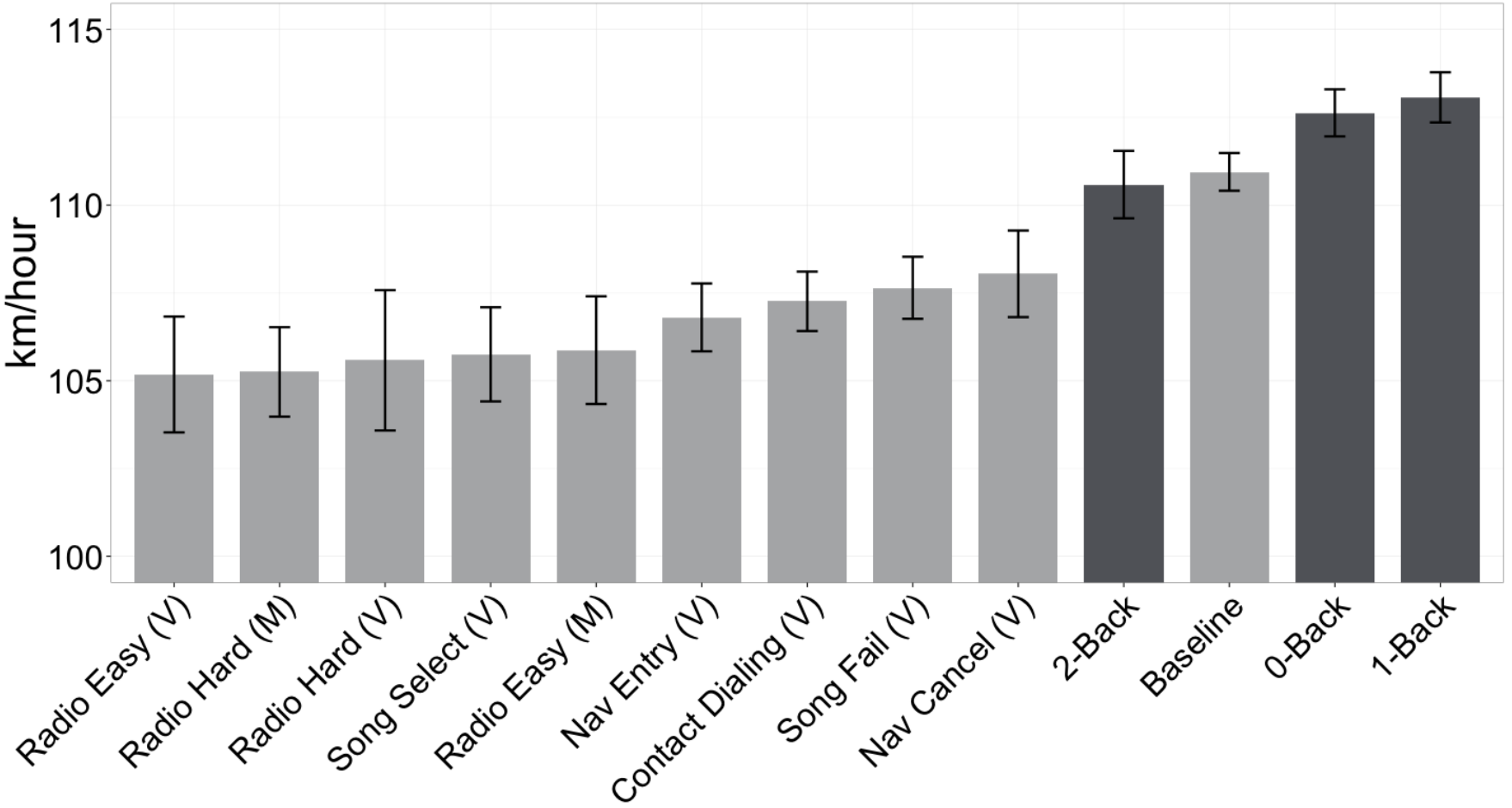
Correspondence > Bryan Reimer, Ph.D. > (617) 452 – 2177 > reimer@mit.edu

Skin Conductance



Correspondence > Bryan Reimer, Ph.D. > (617) 452 – 2177 > reimer@mit.edu

Mean Velocity

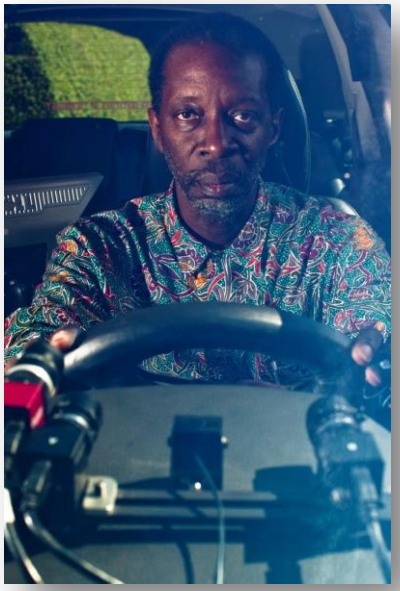


Correspondence > Bryan Reimer, Ph.D. > (617) 452 – 2177 > reimer@mit.edu

Visual Attention

One needs to question the validity of eye tracking data

- Goal -> faceLab data reduced to categorize eyes on road / eyes to task
- Result -> multiple inconsistencies suggest eye tracking data not reliable
- Outcome -> double coded, mediated data forthcoming



Correspondence > Bryan Reimer, Ph.D. > (617) 452 – 2177 > reimer@mit.edu

A Link Between Phone Use and Risky Driving

(Zhao, Mehler, Reimer, D'Ambrosio, Mehler & Coughlin, 2012)

It's clear [from the scientific literature] that cell phones in and of themselves impair the ability to manage the demands of driving, but the fundamental problem may be a broader pattern of behavior of individuals who are willing to pickup the technology.

Correspondence > Bryan Reimer, Ph.D. > (617) 452 – 2177 > reimer@mit.edu

Conclusions on the Demand of Voice Interfaces

Benefits and Cautions

- Cognitive processing demands (as assessed by physiological activation) lower than expectations
  - Largely less than manual radio tuning benchmark
  - Aligned with 1-back cognitive benchmark
- Provide access to higher complexity activity (nav entry) at cognitive demand levels below accepted benchmarks
- Shows high perceived workload and SCL in failure condition
- Clear compensatory slowing to manage workload
- Results encouraging, but visual attention to HMI / eyes-off-road (TBD)

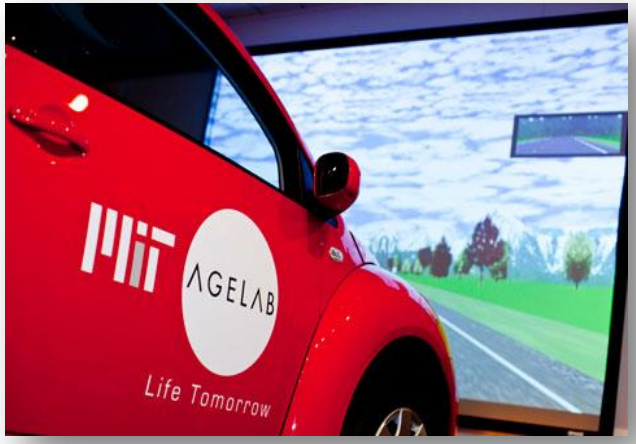
Correspondence > Bryan Reimer, Ph.D. > (617) 452 – 2177 > reimer@mit.edu

Limitations to Generalizability

- Tasks
- Road environment
- Vehicle selection
- Experimental “pacing” of tasks
- “Demand” -> is not clearly linked to safety

Correspondence > Bryan Reimer, Ph.D. > (617) 452 – 2177 > reimer@mit.edu

Questions



Correspondence > Bryan Reimer, Ph.D. > (617) 452 – 2177 > reimer@mit.edu