An Initial Assessment of the Significance of Task Pacing on Self-Report and Physiological Measures of Workload while Driving

Bruce Mehler & Bryan Reimer

The Massachusetts Institute of Technology AgeLab & New England University Transportation Center, Cambridge, MA, USA

E-mail: bmehler@mit.edu

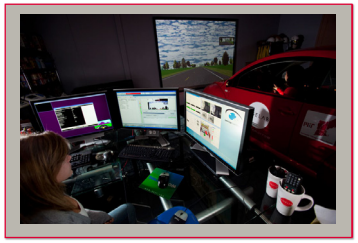
Introduction

Previous work by our group has demonstrated that an auditory presentation – verbal response working memory task (a version of an n-back task) can be successfully employed in simulation and on-road driving contexts to produce graded levels of cognitive demand. This has most clearly been documented in terms of changes in physiological measures of arousal / activation and has led to a number of research groups adopting this n-back task to study the impact of added cognitive demand on various aspects of driving relevant behavior.

The present simulation study consisted of two separate experimental segments (A & B) with the intent of exploring several questions. First, while heart rate and skin conductance level (SCL) have shown statistically significant differences between single task driving and each of 3 levels of the n-back task (0, 1, and 2-back), we had not previously collected self-report ratings of workload in close proximity of the completion of individual tasks. Thus, there is an open question as to what extent self-perceived load compares to the objective physiological measures. Second, working memory load increases in a clearly defined chunks across the 0, 1, and 2-back levels. What is unknown about the task is how much apparent demand is involved in just asking a driver to listen to the stimuli without a requirement to hold them in working memory. This led to testing a "just listen" (blank-back) condition. A final primary question had to do with the extent to which features of the timing of stimuli could be modified to explore aspects of how task pacing impacts an auditory presentation – verbal response interaction in the vehicle.

Methods

Apparatus



The AgeLab simulator running STISIM Drive v.2.08.02 on a 8' by 6' projection screen providing approximately a 40 degree view of a virtual world at 1024 x 768 pixel resolution. Primary task instructions and auditory stimuli are all presented under program control for consistency of presentation.

Participants

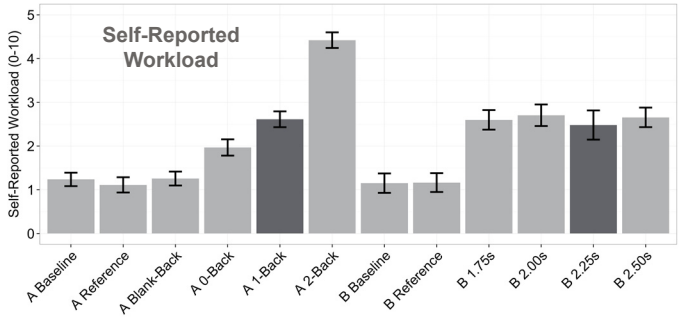
- 38 drivers (23 male) between 20 and 31 years old
- Mean age: 23.4 males (SD=2.8); 24.7 (SD=3.1) females
- License for more than 3 years
- Drive on average 3 or more times per week
- Monetary inducement to balance safe driving and secondary task engagement

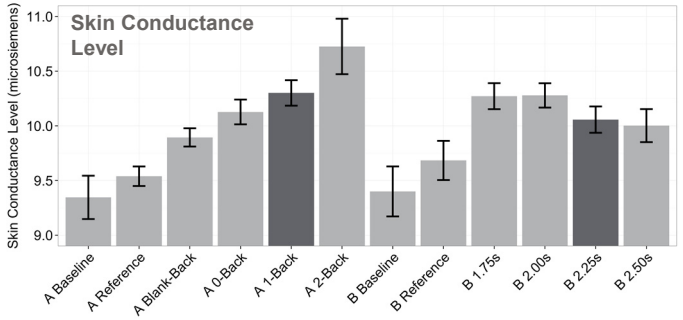
Physiology Measures

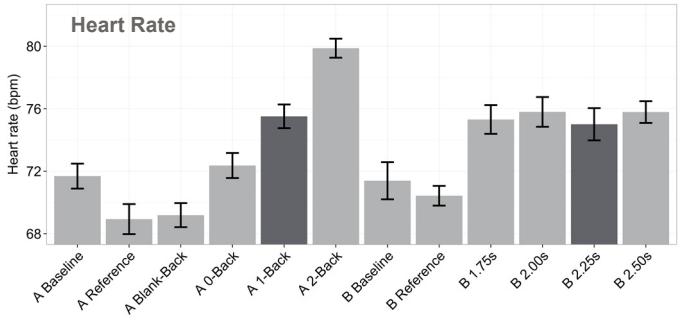
- Modified lead II configuration employed to record EKG at 250 Hz. for heart rate extraction; SCL measured from middle fingers of left hand using thin, dry gold plate sensors to minimize interference with steering wheel

Design

- Simulation consisted of a divided highway with two lanes in each direction, posted speed 65 mph, moderate density of 23 vehicles/mile; participants free to select travel lane and speed
- N-back presentation consisted of single digit numbers 0-9 presented in random order
- Ordering of segments A & B counterbalanced across sample; n-back tasks 30 seconds in duration and randomly ordered within segments
- Segment A included a "just driving" reference, a "just listen" blank-back period, and 0, 1, and 2-backs with standard 2.25 second item spacing
- Segment B presented four 1-back tasks with item presentation intervals of 1.75, 2.0, 2.25, and 2.5 seconds along with a "just driving" reference period







Results

Self-Report - Participants were asked to rate the workload of the task periods following each task on a visual scale with labels of 0-10 (low to high) and mark points that allowed for half number resolution (i.e. 21 points total).

Subjective workload differed significantly between tasks ($X^2_{(11)} = 245.2$, $p < .001$, Friedman test). Sub-tests show significant differences appeared across the ratings of the four levels of the n-back (blank, 0, 1, and 2-back) in Segment A ($X^2_{(3)} = 58.2$, $p < .001$), with subjective workload increasing monotonically with the objective load level ($r = 0.57$, $p < .001$, Pearson correlation). The mean workload ratings of the 1-back tasks in Segment B with different inter-stimulus intervals did not differ in a statistically significant manner ($X^2_{(3)} = 2.59$, $p = .459$).

SCL - Mean skin conductance level (SCL) varied significantly across task periods ($F(11, 407) = 7.77$, $p < .001$, ANOVA with repeated measures). SCL also varied in a statistically significant fashion across the four demand levels of the n-back (blank, 0, 1, and 2-back) in Segment A ($F(3, 111) = 12.8$, $p < .001$) . It did not differ across the four different inter-stimulus timings of the 1-back ($F(3, 111) = 1.3$, $p = .284$) in Segment B. Pairwise comparisons between the blank-back and the non-task periods (baseline driving and the "just drive" reference) were significantly different ($t(37) = 3.34$, $p < .01$; $t(37) = 3.07$, $p < .01$; comparing blank-back to baseline and reference periods, respectively).

Heart Rate - Mean heart rate varied significantly across task type ($F(11, 407) = 20.9$, $p < .001$, ANOVA with repeated measures). Heart rate varied significantly across the four levels of the n-back (blank, 0, 1, and 2-back) in Segment A ($F(3, 111) = 46.1$, $p < .001$) and monotonically with objective demand ($r = .33$, $p < .001$). Mean heart rate did not differ significantly across the different stimulus timings used in Segment B ($F(3, 111) = 0.45$, $p = .717$).

Conclusions

The primary intent in Segment B was to assess the extent to which varying the inter-stimulus interval of a 1-back task provides a useful model for studying one timing aspect of pacing in an auditory presentation – verbal response HMI interface. Looking across self-reported workload ratings, SCL, and heart rate, we are tempted to see a trend suggesting a modest advantage in the 2.25s interval used in earlier studies. This would fit a hypothesis that both speeding-up or drawing-out an interaction from an optimal point increases demand. To the extent that this is true, the effect is subtle as we have tested it here and the fixed memory demand aspects of the n-back may not be ideal for modeling this. Follow-on studies are currently underway or planned to assess other variations in interval duration and a comparison of fixed vs. self-paced responding.

The self-reported workload ratings collected in Segment A establish concurrent validity with physiological data for the use of the 0, 1, and 2-back levels of the n-back task as a method of inducing graded levels of cognitive demand. In addition, heart rate and SCL values indicate that the "blank-back", where a driver is asked to just listen to numbers, also produces a distinguishable physiological response. It is interesting that this is not the case for self-reported workload; drivers clearly do not report that they are being impacted by this "simple" demand. However, heart rate and SCL suggest that there may well be attentional and/or cognitive costs for even this seemingly modest demand on the driver.

Also apparent is a divergence in the direction of SCL and heart rate between baseline and "just drive" reference periods and the "just listen" blank-back condition. This is an excellent example of a "sensory intake" state as described by Lacey et al. where SCL increases and heart rate decreases during attending to stimuli vs. a "sensory rejection" state with active memory load as in the 0, 1, and 2-back. This demonstrates that while increases in heart rate can reasonably be interpreted as evidence for an increase in arousal, the lack of an increase or a drop in heart rate cannot automatically be taken as evidence for its absence. This is one of the reasons why we also collect SCL in research applications.