Cardiac-Activity Measures for Assessing Airport Ramp-Tower Controller’s Workload

Miwa Hayashi, NASA Ames Research Center
Victoria Dulchinos, San Jose State University Foundation

Human Factors and Ergonomics Society (HFES) Annual Meeting
September 19-23, 2016, Washington, DC
Background

- Subjective measures of workload have known shortcomings.
  - “Subjective”
  - Low sensitivity
Low Sensitivity

“Please rate your workload level on the scale of 1 to 7, where 1 is the lowest workload and 7 is the highest.”

This scenario should cause high workload...

© 2016 Miwa Hayashi
Another case of low sensitivity

Start

5 minutes later

5 minutes later

End

1

1

1

1

Really?
Subjective measures of workload have known shortcomings.

- “Subjective”
- Low sensitivity
- Sparse data, likely missing the important event
- Potentially distracting, if measured in real time

Physiological measures may potentially address these weaknesses.

“Objective”
“??”
“Continuous”
“Non-intrusive”
Goal

To examine if mean heart rate (HR) and heart rate variability (HRV) can be used to measure the controller workload in our air-traffic-control simulation evaluation studies.

- Compared mean HR and HRV with the real-time self-reported subjective workload rating results.
- Assessed if mean HR and HRV could replace the subjective measures (e.g., in field tests).
Mean Heart Rate (HR)

- The average number of beats/minute.
  - Derived from the “RR intervals.”

- Considered to reflect an overall level of general arousal, physical work, task demands, and emotional response. (Wierwille & Eggemeier, 1993)
Heart Rate Variability (HRV)

- Measure of variability in the RR intervals.

- Thought to reflect the balance of autonomic nervous system:
  - HF power (0.15 – 0.4 Hz): Parasympathetic activity
  - LF power (0.04 – 0.15 Hz): Parasympathetic and sympathetic activities
  - MF power suppression (0.08 – 0.15 Hz): Increased cognitive effort
Past Studies

- HRV MF suppression used to measure workload:

<table>
<thead>
<tr>
<th>Authors (Year)</th>
<th>Studied Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vicente, Thornton, &amp; Moray (1987)</td>
<td>Low-fidelity hovercraft course-tracking simulation</td>
</tr>
<tr>
<td>Rowe, Sibert, &amp; Irwin (1998)</td>
<td>Air-traffic-control game</td>
</tr>
</tbody>
</table>

- Skeptics:
  - Inconsistent MF results for AGARD-STRESS battery task workload (Nickel & Nachreiner, 2003)
  - Large individual differences in stress reactions of autonomic nervous system (Berntson & Cacioppo, 2004)
Methods
Airport Ramp Tower Simulation

- Simulation evaluation of a NASA’s departure-metering decision-support tool, Spot and Runway Departure Advisor (SARDA).

- 6 Charlotte airport ramp-tower controllers
- 16 runs per controller
- 65-70 minutes departure “push”
- Self-reported subjective workload rating at every 5 minute
- Resulted in 10-12% taxi fuel saving
- No increase in the controller workload
ECG Recording

- RR intervals were recorded with Firstbeat Bodyguard 2 (BG2).
  - Attached to the body via 2 electrodes.
  - Sampling rate = 1000 Hz

- No activity constraint
  - Free to sit, stand, and walk around.
  - OK to smoke or drink tea/coffee during a break.
Computation of Mean HR and HRV

1. Artifacts in the RR interval data were removed. (No replacement.)

2. Mean HR were computed within the 2-minute windows around the sampling times of the real-time workload ratings.

3. Within the same 2-minute windows, MF, HF, and the total power (0.04 – 0.15 Hz) HRV were computed.
   - Lomb-Scargle Periodogram algorithm was used to estimate the power spectral density.

4. MF and HF were normalized with the total power.
Linear Mixed Model (LMM) regression was applied.

\[ Y = WL + P + WL^*P + \epsilon \]

- Mean HR, HRV MF (normalized), or HRV HF (normalized)
- Workload rating
  - Fixed, continuous effect
- Participant
  - Random, categorical effect
- Interaction
- Error
Results & Discussion
Mean HR Results

- The LMM did not find statistical significance in workload (WL) effect.

- The graph shows only a weak trend.
HRV Results

For the normalized MF, WL effect was statistically significant ($p < 0.01$).

However, the estimated coefficient was in the wrong direction (0.015; $SE = 0.006$).

The total power also increased when WL = 3 or 4 (again, the wrong direction).
Discussion

- **Mean HR:** only weak correlation with the workload ratings
  - Subjective measures are not necessarily the true state of workload.

- **HRV-MF and HRV-Total power:** contradictory trend directions
  - Were they more “relaxed” when they reported WL = 3 or 4? (Unlikely.)
  - More plausible explanation: HRV must have sensed something else.
    - Increased speech when traffic volume was high.
    - Posture change, walking around, sipping water, etc.
Summary

- In our airport ramp-tower simulation, we found that mean HR was only weakly correlated with the controller’s self-reported workload levels.

- HRV results were contradictory and inconclusive.

- Until further research is conducted to understand the effects of speech, posture changes, etc., using HR or HRV measures as a sole mean of workload assessment in field tests is not recommended.

- It is recommended to measure HR and HRV along with subjective measures.