

Published online: 1-24-2018

## Physiological and Cognitive Factors Related to Human Performance During the Grand Canyon Rim-to-Rim Hike

Kristin Divis

*Sandia National Laboratories, NM, kmdivis@sandia.gov*

Clifford Anderson-Bergman

*Sandia National Laboratories, CA, ciande@sandia.gov*

Robert Abbott

*Sandia National Laboratories, NM, rgabbot@sandia.gov*

Victoria Newton

*Sandia National Laboratories, NM, vnewton@sandia.gov*

Glory Emmanuel-Aviña

*Sandia National Laboratories, CA, gremman@sandia.gov*

Follow this and additional works at: <https://docs.lib.purdue.edu/jhpee>



Part of the [Cognitive Psychology Commons](#), [Medical Physiology Commons](#), and the [Physiological Processes Commons](#)

---

### Recommended Citation

Divis, Kristin; Anderson-Bergman, Clifford; Abbott, Robert; Newton, Victoria; and Emmanuel-Aviña, Glory (2018) "Physiological and Cognitive Factors Related to Human Performance During the Grand Canyon Rim-to-Rim Hike," *Journal of Human Performance in Extreme Environments*: Vol. 14 : Iss. 1 , Article 5.

DOI: 10.7771/2327-2937.1095

Available at: <https://docs.lib.purdue.edu/jhpee/vol14/iss1/5>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact [epubs@purdue.edu](mailto:epubs@purdue.edu) for additional information.

This is an Open Access journal. This means that it uses a funding model that does not charge readers or their institutions for access. Readers may freely read, download, copy, distribute, print, search, or link to the full texts of articles. This journal is covered under the [CC BY-NC-ND license](#).

---

## Physiological and Cognitive Factors Related to Human Performance During the Grand Canyon Rim-to-Rim Hike

### Cover Page Footnote

Corresponding author: Glory Emmanuel Aviña, Principal Investigator, gremman@sandia.gov Funding for Wearables At The Canyon for Health is provided by the Defense Threat Reduction Agency, Project CB10359.

# Physiological and Cognitive Factors Related to Human Performance During the Grand Canyon Rim-to-Rim Hike

Kristin M. Divis

*Sandia National Laboratories, NM*

Clifford Anderson-Bergman

*Sandia National Laboratories, CA*

Rob Abbott and Victoria E. Newton

*Sandia National Laboratories, NM*

Glory Emmanuel-Aviña

*Sandia National Laboratories, CA*

---

## Abstract

Exposure to extreme environments is both mentally and physically taxing, leading to suboptimal performance and even life-threatening emergencies. Physiological and cognitive monitoring could provide the earliest indicator of performance decline and inform appropriate therapeutic intervention, yet little research has explored the relationship between these markers in strenuous settings. The Rim-to-Rim Wearables at the Canyon for Health (R2R WATCH) study is a research project at Sandia National Laboratories funded by the Defense Threat Reduction Agency to identify which physiological and cognitive phenomena collected by non-invasive wearable devices are the most related to performance in extreme environments. In a pilot study, data were collected from civilians and military warfighters hiking the Rim-to-Rim trail at the Grand Canyon. Each participant wore a set of devices collecting physiological, cognitive, and environmental data such as heart rate, memory, ambient temperature, etc. Promising preliminary results found correlates between physiological markers recorded by the wearable devices and decline in cognitive abilities, although further work is required to refine those measurements. Planned follow-up studies will validate these findings and further explore outstanding questions.

*Keywords:* physiological markers, cognitive markers, human performance, Grand Canyon

---

## Introduction

Warfighters must remain healthy to deliver peak performance and ensure mission success. Adventurous—or just unlucky—civilians are exposed to extreme environments each year, potentially leading to life-threatening situations. Real-time physiological and cognitive monitoring could provide the earliest indication of critically declining performance (e.g., due to fatigue, hyponatremia, exposure to a biological and/or chemical agent, etc.), enabling rapid therapeutic intervention and preventing further performance decline. Similar monitoring is becoming more prevalent in other fields: as Golden State Warriors assistant general manager Kirk Lacob put it, athletes “wear devices that help gauge a player’s fatigue by tracking everything from heart rate to biomechanical load exerted on his legs [because] if Steph Curry is not healthy, we’re not winning” (Leung, 2015).

While the recent explosion in wearable and agile devices presents an opportunity to collect data on various physiological and cognitive performance metrics, it is still unclear which markers are most pertinent and reliable for rapid indication of emerging illness, for determining likelihood of task success, or for determining a cause for a detected health decline. Most research indicative of health deterioration (1) uses laboratory settings or mild tasks to gauge performance and (2) does not examine both physiological and cognitive performance metrics.

The Rim-to-Rim Wearables at the Canyon for Health (R2R WATCH) study at Sandia National Laboratories, funded by the Defense Threat Reduction Agency (Project CB10359), was designed to address these gaps and test the utility of commercial off-the-shelf physiological and cognitive monitoring technologies for their ability to detect early and accurately the

---

Corresponding author: Glory Emmanuel-Aviña, Principal Investigator, Email: gremman@sandia.gov.

first signs of declining health in extreme environments. In partnership with the University of New Mexico Emergency Medicine team, the initial pilot study equipped volunteers with a variety of physiological and cognitive monitoring devices while they hiked the grueling 24.2-mile Rim-to-Rim (R2R) trail at the Grand Canyon (for full study methodology, see Emmanuel-Aviña et al., 2017).

### Physiological and Cognitive Markers of Performance Decline

#### *Physiological*

Physiological correlates of performance decline have been studied in both typical (e.g., 30 minutes of moderate exercise) and atypical (e.g., scaling Mount Everest) environments. The R2R WATCH study was primarily motivated by previous work measuring performance in the latter, where participants were exposed to abnormal environmental stressors and/or required to expend substantial physical effort.

The influence of extremes in external temperature is well studied (for a review, see Rodahl, 2002). Heat stress leads to increased heart rate as the cardiovascular system also works to cool the body by transporting heat from the core to the extremities, where it can dissipate. This increase in heart rate occurs regardless of overall level of physical exertion, and fluctuates in synchronicity with environmental temperature changes (see Rodahl, 2002). Workers at an aluminum production plant saw a 20% increase in cardiovascular workload when exposed to heat stress (Rodahl, 1989); glass bangle workers saw a similar increase and furthermore took longer to return to a normal heart rate once removed from the hot environment (Rastogi, Gupta, Husain, & Mathur, 1990).

Dehydration, leading to hypohydration, affects cardiovascular performance and the body's ability to properly regulate temperature. As Sawka, Montain, and Latzka (2001) report in a review of the field, core body temperature rises on average 0.1–0.25°C for every percent of dehydration-induced body weight loss, with even larger increases seen when physically exerting oneself in hot conditions (Adolph et al., 1947; Montain & Coyle, 1992; Sawka, Young, Francesconi, Muza, & Pandolf, 1985; Strydom & Holdsworth, 1968). Being fit and accustomed to hot environments normally lends one a performance advantage; hypohydration (e.g., 5% body weight reduced) negates that advantage (Buskirk, Iampietro, & Bass, 1958; Cadarette, Sawka, Toner, & Pandolf, 1984; Sawka, Toner, Francesconi, & Pandolf, 1983). Under heat stress and hypohydration, overall cardiac output is decreased as heart rate increases but stroke volume decreases (see Sawka et al., 2001). In general, hypohydration reduces athletic performance, with greater losses seen in tasks requiring more endurance (e.g., 5% reduction for a 10,000-meter race compared to 3% for a 1,500-meter race; Armstrong, Costill, & Fink, 1985).

The effects of extreme fatigue induced via sleep deprivation were studied in a sample of mountain ultra-marathon

runners (Poussel et al., 2015). Runners who adopted a sleep-management strategy of increasing sleep time a few days before the race finished faster on average. One interpretation of these findings is that runners who begin to show signs of fatigue are revealing early signs of performance decrement. Runners with higher reported levels of drowsiness also took longer to complete the race.

Veltman and Gaillard (1998) linked physiological measures and workload during a challenging flight simulation. Heart rate, blood pressure (from beat to beat), respiration, and eye blinks were recorded as participants performed complex flight and memory tasks. All measures were sensitive to large changes in workload; heart period (a combination of heart rate variability and blood pressure variability less influenced by respiration) was the most sensitive to relatively small changes in level of task difficulty. A study of mountaineers scaling the Cho-Oyo in Tibet examined blood pressure, pulse, skin resistance, blood pressure relaxation, and anxiety (Stück, Balzer, Hecht, & Schröder, 2005). The participants progressed from inhibition of overload to hypersensitivity to exhaustion. Notably, the psychophysiological measures predicted decline prior to the alpinists' awareness of that decline.

The physiological markers reported in these studies highlight how physiological monitoring may be used to more quickly assess and alleviate health risks and performance decline in strenuous environments.

#### *Cognitive*

A substantial body of literature supports the claim that extreme fatigue and stress on the body, induced by the physical environment, have negative effects on cognitive functioning. Past research indicates that even mild thermal stress may affect human performance (Enander, 1989; Hancock & Vasmatazidis, 1998). Hocking, Stilberstain, Lau, Stough, and Roberts (2001) showed an association between extreme temperatures and deficits in working memory and information processing as measured by the digit span task and AX-continuous performance task, which is a measure of attention, memory, verbal learning, information processing, and concentration. When looking at the stress–performance relationship, Grether (1973) demonstrated that response time and vigilance tend to share a curvilinear relationship with temperature. Performance increases up to 30°C, at which point it reliably decreases. A series of studies have examined the effects of cold on physical and cognitive performance. Exposure to cold air resulted in decreased performance on serial choice–reaction time tasks (Ellis, 1982; Ellis, Wilcock, & Zaman, 1985) and working memory deficits have been reported after core body temperatures dip beneath 37°C. A study of naval special operations forces during actual winter warfare training found that cold temperatures were associated with decrements in hand strength and fine motor skills (Hyde, Thomas, Schrot, & Taylor, 1997). Additionally, performance

was especially affected when temperature varied over time and for extremely high temperatures (Enander, 1989).

Climbers who completed perceptual, cognitive, and sensory-motor tasks while scaling Mount Denali showed poorer memory and learning and overall slower performance relative to a matched control group (Kramer, Coyne, & Strayer, 1993). Cognitive deficits—particularly in memory—have consistently been associated with altitude change (Muza, Kaminsky, Fulco, Banderet, & Cymerman, 2004). The Spaceflight Cognitive Assessment Tool for Windows (WinsCAT) is a cognitive test battery designed to assess the neurocognitive status of astronauts on missions of long duration at various altitudes (Lowe et al., 2007). Decreased performance in the running memory task of the WinsCAT was reported between 0.5 and 4 hours after ascent; however, similar deficits were no longer present at tests given 12 and 24 hours after ascent. This may indicate habituation to altitude change. Alternatively, it may indicate that cognitive performance is affected by variability in altitude over short periods of time. Additionally, the cognitive deficits reported largely occurred before physiological symptoms of mountain sickness were reported, highlighting the potential use of cognitive markers as early warning signs of decline.

Research has also found that memory, accuracy, reaction time, attention, and cognitive executive functions are impaired by fatigue and stress (Bourne & Yaroush, 2003; Karatsoreos & McEwen, 2010). One study used a computerized cognitive test battery specifically designed for the high-performing astronaut population. The test measured various cognitive domains, including emotional processing, spatial orientation, and risk decision-making. Fatigue, as measured by acute sleep deprivation, was found to negatively affect vigilant attention, cognitive throughput, and abstract reasoning (Basner et al., 2015). Meta-analyses of fatigue and performance literature report consistent findings that fatigue negatively impacts several functions, but in particular visual attention, vigilance, decision-making, and other complex cognitive functions (Bourne & Yaroush, 2003). In addition, simple tasks like drinking water may have extreme consequences if not completed properly, especially in extreme environments such as those encountered during the Grand Canyon R2R hike (Wickens, Keller, & Shaw, 2015).

### *Grand Canyon R2R Environment*

The Grand Canyon 24.2-mile R2R hike represents a rigorous performance task involving an elevation change of nearly 7000 feet from rim to canyon floor and temperature differentials of up to 50°F. While the park service highly discourages tackling the entire R2R hike in a single day, thousands of hikers attempt it each season. The R2R is a rigorous hike, requiring the body to endure fatigue and stress while adapting to rapidly changing environmental conditions; each year, over 250 people are airlifted from the

canyon, many with symptoms of hyponatremia and heat stroke (Garigan & Ristedt, 1999; Ghiglieri & Myers, 2001).

### *Current Study*

The R2R WATCH study was designed to collect, analyze, and link data on physiological, cognitive, and biological markers in order to more quickly and accurately predict performance decline and health risks in extreme environments. The pilot study and initial findings reported in this paper are the first step toward meeting those goals. We collected physiological and cognitive data from civilians and military warfighters attempting to hike the Grand Canyon R2R in a single day. Follow-up studies to validate and expand the results of this pilot study are currently underway. The analyses reported below focus on the pilot cognitive and physiological data overall; future work will pull in the biological data and tease apart potential differences between military and civilian populations.

## **Method**

### *Participants*

Participants included both civilian and military populations. Research assistants identified civilian R2R hikers at the South Kaibab trailhead and asked if they would like to participate in the study. Military participants who were interested in completing the R2R hike were informed of the data collection dates and invited to volunteer for the study.

There were three tiers to the R2R WATCH study (see Emmanuel-Aviña et al., 2017), but this article focuses on the data collected through wearable devices. Over 950 civilian hikers attempted the R2R during our 48-hour data collection period; 38 agreed to participate in the wearables portion of the study (19 males; age in years: *mean* = 46.29, *stdev* = 11.89). Twelve warfighters from a special population in the military also participated in the study (6 males; age in years: *mean* = 35.92, *stdev* = 6.29), leading to 50 total participants. Three subjects chose to run the R2R2R (i.e., South Rim to North Rim and back to South Rim). Data were recorded for these three subjects but excluded from data analyses due to extreme differences in activity.

### **Design and Materials**

Participants were outfitted with one of four packages of devices that they wore while completing the R2R hike. The devices collected physiological, cognitive, location, and environmental data.<sup>1</sup> There were two types of device

<sup>1</sup>This pilot study was part of a larger study of R2R hikers conducted in conjunction with the University of New Mexico Emergency Medicine team. The other two components of the project were pre-hike, mid-hike, and post-hike surveys (288 participants) and blood samples (49 participants). See Emmanuel-Aviña et al. (2017).

Table 1  
Device options for each type of package.

| Device                            | Metrics                                                               | Package |         |                  |                  |
|-----------------------------------|-----------------------------------------------------------------------|---------|---------|------------------|------------------|
|                                   |                                                                       | Basic 1 | Basic 2 | Advanced 1       | Advanced 2       |
| Apple iPod touch (6th generation) | BrainBaseline cognitive assessments                                   | Yes     | Yes     | Yes              | Yes              |
| Fitbit Charge HR                  | Wrist-based optical heart rate monitor; accelerometer; altimeter      |         | Yes     |                  |                  |
| Garmin eTrex 10                   | GPS                                                                   |         | Yes     |                  |                  |
| Garmin fēnix 3 HR                 | Wrist-based optical heart rate monitor; accelerometer; altimeter; GPS |         |         | Yes              |                  |
| Garmin tempe                      | Body temperature (under chest strap)                                  |         |         | Yes              |                  |
| Garmin vívoactive HR              | Wrist-based optical heart rate monitor; accelerometer; altimeter; GPS | Yes     |         |                  |                  |
| LifeBEAM SmartHat                 | Forehead-based optical heart rate monitor                             |         |         | Yes              | Yes              |
| Myontec Mbody Shorts              | Quadri- and hamstring muscle group monitoring                         |         |         | Yes <sup>a</sup> |                  |
| SensorPush                        | Thermometer; hygrometer                                               | Yes     | Yes     | Yes <sup>b</sup> | Yes <sup>b</sup> |
| Suunto Smart Sensor               | ECG heart rate monitor                                                |         |         |                  | Yes              |
| Suunto Spartan Ultra              | Accelerometer; altimeter; GPS                                         |         |         |                  | Yes              |
| Wahoo TICKRx                      | ECG heart rate monitor; accelerometer                                 |         |         | Yes              |                  |

<sup>a</sup>Only included with 5 packages.

<sup>b</sup>Two SensorPushes (one in indirect sunlight as in basic packages; one in direct sunlight).

packages: advanced and basic, each with a preferred option and a secondary option to increase diversity of devices (leading to a total of four package options). Civilian participants wore the basic packages; military participants used the advanced packages. See Table 1 for a list of all packages, devices, and metrics. At the very least, all packages measured location, ambient temperature and humidity, heart rate, cadence, perceived fatigue, and cognitive abilities.

Cognitive assessments were administered on a mobile device using a customized version of Digital Artefact's BrainBaseline application. The cognitive battery included a fatigue questionnaire and three cognitive tasks: visual short-term memory (VSTM; Cowan, 2001; Luck & Vogel, 1997), flanker (Eriksen & Eriksen, 1974; Fan, McCandliss, Fossella, Flombaum, & Posner, 2005), and go/no-go (Conners & Sitarenios, 2011). The fatigue questionnaire queried how mentally fatigued and physically fatigued the participant felt on a scale from 1 to 6. The VSTM task consisted of 50 trials (50% match); the flanker task consisted of 100 trials (50% congruent); and the go/no-go task consisted of 50 trials, with 20% no-go and a delay ranging from 500 to 1850 ms.<sup>2</sup> All tasks were implemented via touch screen buttons on an Apple iPod touch (6th generation).

Data were collected from 50 sets of packages: 32 Basic-1 (preferred) packages, 6 Basic-2 (secondary) packages, 10 Advanced-1 (preferred) packages, and 2 Advanced-2 (secondary) packages.

<sup>2</sup> See <https://www.brainbaseline.com> for additional details on the VSTM, flanker, and go/no-go tasks.

## Procedure

At the South Kaibab trailhead, civilian R2R hikers were briefed on the wearable devices and given the option to participate. Those who consented were outfitted with a fitness watch and GPS tracker, had a SensorPush zip-tied to their bag, and then worked through an initial session of the cognitive battery on the iPod prior to beginning the hike. Military R2R hikers went through a similar procedure except they were given additional devices to wear.

Since the fitness and environment devices passively collect data once turned on, participants were asked to leave them alone unless they needed to adjust fit. The cognitive assessment was to be performed at the beginning and end of the hike and then approximately every three hours in between during natural breaks in the hike. The BrainBaseline application alerted participants with a tone when it had been three hours since finishing the last assessment.

Upon completion of the hike, researchers met hikers at the North Kaibab trailhead, asked them to complete the cognitive assessment one more time, reclaimed the devices, and debriefed the participants.

## Statistical Methodology

Two models were built for predicting decline in cognitive abilities as measured by the BrainBaseline tests. The first model used measures of fatigue that were heavily dependent on the structured nature of the activity. This model was intended to validate that fatigue influenced the cognitive

Table 2

Response time effects for the cognitive battery as a function of proportion up and down the canyon.

| Estimated Effects on Response Time   |          |               |         |
|--------------------------------------|----------|---------------|---------|
|                                      | Estimate | 95% CI        | P value |
| Flanker-congruent: proportion up     | 72.1     | (23.7, 120.6) | 0.004   |
| Flanker-congruent: proportion down   | -8.0     | (-69.1, 53.1) | 0.797   |
| Flanker-incongruent: proportion up   | 49.8     | (-8.2, 107.9) | 0.093   |
| Flanker-incongruent: proportion down | 6.5      | (-67.7, 80.8) | 0.863   |
| Go/no-go: proportion up              | 27.7     | (-23.6, 78.9) | 0.290   |
| Go/no-go: proportion down            | 32.9     | (-22.7, 88.5) | 0.247   |
| VSTM: proportion up                  | 206.8    | (10.8, 402.9) | 0.039   |
| VSTM: proportion down                | 225.2    | (-33, 183.4)  | 0.087   |

Table 3

Accuracy effects for the cognitive battery as a function of proportion up and down the canyon.

| Estimated Effects on Accuracy |          |                  |         |
|-------------------------------|----------|------------------|---------|
|                               | Estimate | 95% CI           | P value |
| Flanker: proportion up        | -0.003   | (-0.052, .046)   | 0.9102  |
| Flanker: proportion down      | -0.002   | (-0.067, 0.063)  | 0.9562  |
| Go/no-go: proportion up       | -0.047   | (-0.088, -0.007) | 0.0229  |
| Go/no-go: proportion down     | -0.030   | (-0.074, 0.015)  | 0.1878  |
| VSTM: proportion up           | -0.124   | (-0.184, -0.064) | 0.0001  |
| VSTM: proportion down         | -0.098   | (-0.178, -0.019) | 0.0155  |

Table 4

Response time effects for the cognitive battery as a function of heart rate (HR) zone.

| Estimated Effects on Response Time          |          |               |         |
|---------------------------------------------|----------|---------------|---------|
|                                             | Estimate | 95% CI        | P value |
| Flanker-congruent: HR below 120             | 2.4      | (-23.9, 28.7) | 0.859   |
| Flanker-congruent: HR between 120 and 160   | 23.7     | (1.1, 46.3)   | 0.040   |
| Flanker-congruent: HR above 160             | -7.5     | (-23.2, 8.1)  | 0.345   |
| Flanker-incongruent: HR below 120           | 16.4     | (-16, 48.7)   | 0.321   |
| Flanker-incongruent: HR between 120 and 160 | 27.0     | (-0.7, 54.6)  | 0.056   |
| Flanker-incongruent: HR above 160           | -1.3     | (-21.9, 19.3) | 0.902   |
| Go/no-go: HR below 120                      | -1.7     | (-27.3, 23.8) | 0.894   |
| Go/no-go: HR between 120 and 160            | 20.9     | (-0.9, 42.7)  | 0.061   |
| Go/no-go: HR above 160                      | 0.6      | (-14, 15.1)   | 0.938   |
| VSTM: HR below 120                          | -46.1    | (-147, 54.8)  | 0.370   |
| VSTM: HR between 120 and 160                | 104.1    | (17.8, 190.5) | 0.018   |
| VSTM: HR above 160                          | 10.6     | (-54, 75.3)   | 0.748   |

measurements; it would not generalize to a non-structured activity. The second model used noisier measures of fatigue that could be applied to an unstructured environment.

For each model, a summary statistics analysis was performed by extracting the time of test for each BrainBaseline examination.<sup>3</sup> Using this test time, a variety of summary

<sup>3</sup>Note that while participants were encouraged to complete the cognitive battery approximately every three hours, they varied in their compliance with those instructions and how quickly they made it across the canyon. See Figure 1 in the appendix for a representation of where each test (and how many) were taken. One benefit of this variance across participants is that it helps us to pull out learning effects (e.g., getting better at the cognitive battery after each attempt).

statistics designed to capture previous levels of activity based on the collected device data was built. These summary statistics were then regressed on performance on the cognitive tasks using a linear mixed effects model to account for repeated measures. Indicator variables for test number (i.e., first test, second test, etc.) were included to capture the learning effect of subjects in both models. Response variables were performance on the flanker, go/no-go, and VSTM tasks. Each of these tests had two responses measures: accuracy and time between prompt and user's response. The flanker task was divided into congruent and incongruent trials. A separate regression model was fit for each response value.

In the first model, fatigue was captured using location to determine progress of subjects at time of the cognitive battery. Two variables were constructed: percent descended down the canyon and percent ascended back up, as measured in elevation change. *A priori*, it was believed that descending would invoke light fatigue (and thus light effect

on cognitive abilities) and ascending would invoke heavier fatigue.

While the methods in the first model provide a robust measure of relative fatigue, they do not provide a method that could be generally applicable in an unstructured environment. In the second model, this was addressed by

Table 5  
Accuracy effects for the cognitive battery as a function of heart rate (HR) zone.

| Estimated Effects on Accuracy    |          |                 |         |
|----------------------------------|----------|-----------------|---------|
|                                  | Estimate | 95% CI          | P value |
| Flanker: HR below 120            | -0.008   | (-0.035, 0.019) | 0.556   |
| Flanker: HR between 120 and 160  | 0.002    | (-0.02, 0.025)  | 0.829   |
| Flanker: HR above 160            | -0.003   | (-0.021, 0.015) | 0.736   |
| Go/no-go: HR below 120           | -0.004   | (-0.025, 0.016) | 0.688   |
| Go/no-go: HR between 120 and 160 | 0.002    | (-0.016, 0.019) | 0.852   |
| Go/no-go: HR above 160           | -0.002   | (-0.014, 0.01)  | 0.738   |
| VSTM: HR below 120               | -0.009   | (-0.042, 0.025) | 0.616   |
| VSTM: HR between 120 and 160     | -0.024   | (-0.052, 0.005) | 0.104   |
| VSTM: HR above 160               | 0.004    | (-0.017, 0.026) | 0.690   |

Table 6  
Response time effects for the cognitive battery as a function of proportion up and down the canyon, including learning effects.

| Estimated Effects on Response Time   |          |                   |         |
|--------------------------------------|----------|-------------------|---------|
|                                      | Estimate | 95% CI            | P value |
| Flanker congruent: intercept         | 571.1    | (544.8, 597.4)    | 0       |
| Flanker congruent: proportion up     | 72.1     | (23.7, 120.6)     | 0.004   |
| Flanker congruent: proportion down   | -8       | (-69.1, 53.1)     | 0.797   |
| Flanker congruent: test number 2     | -53.9    | (-115.5, 7.6)     | 0.086   |
| Flanker congruent: test number 3     | -64.3    | (-135.6, 6.9)     | 0.077   |
| Flanker congruent: test number 4     | -104.9   | (-190.6, -19.2)   | 0.016   |
| Flanker congruent: test number 5     | -117.8   | (-215.3, -20.3)   | 0.018   |
| Flanker congruent: test number 6     | -154.6   | (-285.2, -24.1)   | 0.02    |
| Flanker incongruent: intercept       | 625.3    | (594.8, 655.8)    | 0       |
| Flanker incongruent: proportion up   | 49.8     | (-8.2, 107.9)     | 0.093   |
| Flanker incongruent: proportion down | 6.5      | (-67.7, 80.8)     | 0.863   |
| Flanker incongruent: test number 2   | -87.6    | (-162.1, -13.1)   | 0.021   |
| Flanker incongruent: test number 3   | -106.1   | (-191.3, -20.8)   | 0.015   |
| Flanker incongruent: test number 4   | -122.8   | (-224.4, -21.3)   | 0.018   |
| Flanker incongruent: test number 5   | -120.4   | (-235.8, -5)      | 0.041   |
| Flanker incongruent: test number 6   | -171.2   | (-326.1, -16.2)   | 0.03    |
| Go/no-go: intercept                  | 549.4    | (522.7, 576)      | 0       |
| Go/no-go: proportion up              | 27.7     | (-23.6, 78.9)     | 0.29    |
| Go/no-go: proportion down            | 32.9     | (-22.7, 88.5)     | 0.247   |
| Go/no-go: test number 2              | -74      | (-129.7, -18.3)   | 0.009   |
| Go/no-go: test number 3              | -95.1    | (-161.1, -29.1)   | 0.005   |
| Go/no-go: test number 4              | -105.2   | (-186.2, -24.1)   | 0.011   |
| Go/no-go: test number 5              | -97.6    | (-190.8, -4.4)    | 0.04    |
| Go/no-go: test number 6              | -121.8   | (-248.9, 5.2)     | 0.06    |
| VSTM: intercept                      | 1139.4   | (1029.3, 1249.5)  | 0       |
| VSTM: proportion up                  | 206.8    | (10.8, 402.9)     | 0.039   |
| VSTM: proportion down                | 225.2    | (-33, 483.4)      | 0.087   |
| VSTM: test number 2                  | -405.8   | (-667.2, -144.4)  | 0.002   |
| VSTM: test number 3                  | -463.6   | (-763.3, -163.9)  | 0.002   |
| VSTM: test number 4                  | -491.8   | (-846.6, -137)    | 0.007   |
| VSTM: test number 5                  | -503.8   | (-911.3, -96.4)   | 0.015   |
| VSTM: test number 6                  | -693.2   | (-1183.8, -202.6) | 0.006   |



using only physiological measurements. To do this, three new variables were constructed: number of hours with heart rate at 0–120 beats per minute (bpm), number of hours at 120–160 bpm, and number of hours at 160+ bpm. These three categories are a reduction of the five standard heart rate zones (Borreson & Lambert, 2009). This captures amount of time spent on light, moderate, and heavy activity. Again, *a priori*, it was believed that exposure to light activity should have little effect on cognitive performance, while more invigorating activity should lead to decline.

Data quality was an issue for the heart rate measures. Several of the devices reported clearly inaccurate data (large amounts of missing data, sustained heart rates above 200 bpm, etc.). Records with clearly inaccurate data (as determined by visual examination of the times series) were dropped from the data set. A total of 13 devices' heart rate data were dropped, although several of these devices belonged to subjects with multiple devices recording heart rate. In total, four subjects (of 47) were excluded from the heart analysis due to a lack of reliable data. In general, it was found that the chest-based EKG devices were more reliable than the optical devices.

## Results

Tables 2 and 3 show the main results from the first model, including the estimated effects of fatigue, as captured in proportion down and back up the canyon. In this

model, the estimated effects were the changes in response (either response time in milliseconds or accuracy) as proportion up/down the canyon increased. Tables 6 and 7 in the appendix also show estimated learning effects.

We hypothesized that fatigue would have a positive effect on response time (i.e., increase in response time). Fatigue was hypothesized to have a negative effect on accuracy. In 13/14 estimated effects, this trend was observed (*p*-value from sign test: 0.0009). Note that not all of the individual effects were statistically significant. In particular, the estimated effect on accuracy of the flanker test was extraordinarily low. *Post hoc* inspection revealed that the baseline accuracy for the flanker task was very high (94%), suggesting subjects performed nearly perfectly in all conditions. This highlights that accuracy on the flanker task is not useful as a response variable for this or future studies, but it should be noted that *response time* for the flanker task was quite responsive to fatigue. Whether response time in the flanker task provides information beyond response time metrics obtained from other tasks has yet to be analyzed.

Tables 4 and 5 show the results of the second model using heart rate as an indicator of fatigue. Tables 8 and 9 in the appendix also show learning effect estimates.

There were no significant effects observed on accuracy across all the variables. A consistent positive trend for heart rate between 120 and 160 bpm was observed for response time, with two of four estimated effects being statistically

Table 7

Accuracy effects for the cognitive battery as a function of proportion up and down the canyon, including learning effects.

| Estimated Effects on Accuracy |          |                  |                    |
|-------------------------------|----------|------------------|--------------------|
|                               | Estimate | 95% CI           | <i>P</i> value     |
| Flanker: intercept            | 0.942    | (0.917, 0.967)   | 0                  |
| Flanker: proportion up        | −0.003   | (−0.052, 0.046)  | 0.9102             |
| Flanker: proportion down      | −0.002   | (−0.067, 0.063)  | 0.9562             |
| Flanker: test number 2        | 0.026    | (−0.038, 0.091)  | 0.4252             |
| Flanker: test number 3        | 0.037    | (−0.034, 0.109)  | 0.3065             |
| Flanker: test number 4        | 0.03     | (−0.053, 0.113)  | 0.4799             |
| Flanker: test number 5        | 0.039    | (−0.055, 0.133)  | 0.4193             |
| Flanker: test number 6        | 0.041    | (−0.087, 0.168)  | 0.5315             |
| Go/no-go: intercept           | 0.959    | (0.939, 0.979)   | 0                  |
| Go/no-go: proportion up       | −0.047   | (−0.088, −0.007) | 0.0229             |
| Go/no-go: proportion down     | −0.03    | (−0.074, 0.015)  | 0.1878             |
| Go/no-go: test number 2       | 0.026    | (−0.018, 0.071)  | 0.2451             |
| Go/no-go: test number 3       | 0.045    | (−0.006, 0.097)  | 0.0863             |
| Go/no-go: test number 4       | 0.072    | (0.009, 0.135)   | 0.0255             |
| Go/no-go: test number 5       | 0.071    | (−0.001, 0.143)  | 0.0535             |
| Go/no-go: test number 6       | 0.078    | (−0.021, 0.177)  | 0.1215             |
| VSTM: intercept               | 0.734    | (0.702, 0.766)   | 0                  |
| VSTM: proportion up           | −0.124   | (−0.184, −0.064) | $1 \times 10^{-4}$ |
| VSTM: proportion down         | −0.098   | (−0.178, −0.019) | 0.0155             |
| VSTM: test number 2           | 0.086    | (0.005, 0.166)   | 0.0369             |
| VSTM: test number 3           | 0.149    | (0.057, 0.24)    | 0.0015             |
| VSTM: test number 4           | 0.15     | (0.042, 0.258)   | 0.0065             |
| VSTM: test number 5           | 0.209    | (0.085, 0.333)   | $9 \times 10^{-4}$ |
| VSTM: test number 6           | 0.173    | (0.023, 0.322)   | 0.0233             |

Table 8

Response time effects for the cognitive battery as a function of heart rate (HR) zone, including learning effects.

| Estimated Effects on Response Time          |          |                  |         |
|---------------------------------------------|----------|------------------|---------|
|                                             | Estimate | 95% CI           | P value |
| Go/no-go: intercept                         | 559.5    | (530.9, 588.2)   | 0       |
| Go/no-go: test number 2                     | -61.2    | (-102.5, -19.9)  | 0.004   |
| Go/no-go: test number 3                     | -80.9    | (-144.4, -17.3)  | 0.013   |
| Go/no-go: test number 4                     | -67.6    | (-152.4, 17.3)   | 0.119   |
| Go/no-go: test number 5                     | -84.7    | (-187.3, 17.9)   | 0.106   |
| Go/no-go: test number 6                     | -126.2   | (-277.6, 25.1)   | 0.102   |
| Go/no-go: HR below 120                      | -1.7     | (-27.3, 23.8)    | 0.894   |
| Go/no-go: HR between 120 and 160            | 20.9     | (-0.9, 42.7)     | 0.061   |
| Go/no-go: HR above 160                      | 0.6      | (-14, 15.1)      | 0.938   |
| Flanker congruent: intercept                | 575.6    | (545.3, 605.9)   | 0       |
| Flanker congruent: test number 2            | -73.2    | (-115.5, -30.8)  | 0.001   |
| Flanker congruent: test number 3            | -82.8    | (-146.5, -19.2)  | 0.011   |
| Flanker congruent: test number 4            | -85.1    | (-170.4, 0.1)    | 0.05    |
| Flanker congruent: test number 5            | -103.5   | (-207, 0)        | 0.05    |
| Flanker congruent: test number 6            | -170.2   | (-323.3, -17)    | 0.029   |
| Flanker congruent: HR below 120             | 2.4      | (-23.9, 28.7)    | 0.859   |
| Flanker congruent: HR between 120 and 160   | 23.7     | (1.1, 46.3)      | 0.04    |
| Flanker congruent: HR above 160             | -7.5     | (-23.2, 8.1)     | 0.345   |
| Flanker incongruent: intercept              | 628.6    | (593.7, 663.6)   | 0       |
| Flanker incongruent: test number 2          | -104.5   | (-157.5, -51.4)  | 0       |
| Flanker incongruent: test number 3          | -142.6   | (-219.8, -65.3)  | 0       |
| Flanker incongruent: test number 4          | -140.4   | (-243.3, -37.5)  | 0.007   |
| Flanker incongruent: test number 5          | -156.1   | (-281.9, -30.4)  | 0.015   |
| Flanker incongruent: test number 6          | -247.3   | (-432.4, -62.1)  | 0.009   |
| Flanker incongruent: HR below 120           | 16.4     | (-16, 48.7)      | 0.321   |
| Flanker incongruent: HR between 120 and 160 | 27       | (-0.7, 54.6)     | 0.056   |
| Flanker incongruent: HR above 160           | -1.3     | (-21.9, 19.3)    | 0.902   |
| VSTM: intercept                             | 1147.6   | (1030.8, 1264.4) | 0       |
| VSTM: test number 2                         | -194.9   | (-359.2, -30.5)  | 0.02    |
| VSTM: test number 3                         | -202.5   | (-442.7, 37.8)   | 0.099   |
| VSTM: test number 4                         | -170.7   | (-482.5, 141)    | 0.283   |
| VSTM: test number 5                         | -97.8    | (-482.7, 287.1)  | 0.618   |
| VSTM: test number 6                         | -408.5   | (-912.5, 95.5)   | 0.112   |
| VSTM: HR below 120                          | -46.1    | (-147, 54.8)     | 0.37    |
| VSTM: HR between 120 and 160                | 104.1    | (17.8, 190.5)    | 0.018   |
| VSTM: HR above 160                          | 10.6     | (-54, 75.3)      | 0.748   |

significant. No statistically significant effects were found for heart rate below 120 bpm and heart rate above 160 bpm. *Post hoc*, it was hypothesized that the lack of effect found for heart rate over 160 bpm could be explained for two reasons. First, there was very little data collected in this range; only 7 subjects achieved heart rates over 160 bpm for more than 30 minutes. Second, this represented significant physical effort. Subjects that chose to exert this level of effort were likely to be unusually fit and thus less affected by fatigue.

## Discussion

A major goal of this study was to examine whether physiological data collected from wearable devices could be linked to decline in cognitive abilities. Initial findings demonstrated that various fatigue measurements captured in the device data were correlated with reduction in

cognitive abilities, suggesting decline in cognitive abilities could be predicted by measurements collected by wearable devices.

Overall we found significant relationships between physiological data such as heart rate and cognitive ability, as measured by the flanker, go/no-go, and VSTM tasks. This opens the door to identify other early health indicators of performance that are currently not available on wearable devices.

While initial analyses revealed these simple correlations, building a model to precisely predict reduction in cognitive abilities will require more sophisticated techniques than those presented here. We suggest two areas of potential improvement. First is to improve the quality of the data collected, in regard to both device data and cognitive measurements. Several subjects were dropped from the heart rate analysis due to clearly degenerate data. Others with questionable measures were included, thus adding noise to the covariates. Collecting data in extreme environments

Table 9

Accuracy effects for the cognitive battery as a function of heart rate (HR) zone, including learning effects.

| Estimated Effects on Accuracy    |          |                 |         |
|----------------------------------|----------|-----------------|---------|
|                                  | Estimate | 95% CI          | P value |
| Go/no-go: intercept              | 0.958    | (0.937, 0.98)   | 0       |
| Go/no-go: test number 2          | 0.009    | (−0.025, 0.043) | 0.6066  |
| Go/no-go: test number 3          | 0.002    | (−0.049, 0.053) | 0.937   |
| Go/no-go: test number 4          | 0.016    | (−0.051, 0.084) | 0.6378  |
| Go/no-go: test number 5          |          | (−0.073, 0.091) | 0.8255  |
| Go/no-go: test number 6          | −0.002   | (−0.123, 0.118) | 0.9702  |
| Go/no-go: HR below 120           | −0.004   | (−0.025, 0.016) | 0.6883  |
| Go/no-go: HR between 120 and 160 | 0.002    | (−0.016, 0.019) | 0.8516  |
| Go/no-go: HR above 160           | −0.002   | (−0.014, 0.01)  | 0.7378  |
| Flanker: intercept               | 0.943    | (0.915, 0.972)  | 0       |
| Flanker: test number 2           | 0.031    | (−0.014, 0.076) | 0.1742  |
| Flanker: test number 3           | 0.043    | (−0.021, 0.106) | 0.185   |
| Flanker: test number 4           | 0.042    | (−0.042, 0.126) | 0.3263  |
| Flanker: test number 5           | 0.05     | (−0.053, 0.153) | 0.3407  |
| Flanker: test number 6           | 0.049    | (−0.102, 0.2)   | 0.5236  |
| Flanker: HR below 120            | −0.008   | (−0.035, 0.019) | 0.556   |
| Flanker: HR between 120 and 160  | 0.002    | (−0.02, 0.025)  | 0.8294  |
| Flanker: HR above 160            | −0.003   | (−0.021, 0.015) | 0.7357  |
| VSTM: intercept                  | 0.743    | (0.704, 0.781)  | 0       |
| VSTM: test number 2              | 0.011    | (−0.043, 0.065) | 0.6933  |
| VSTM: test number 3              | 0.041    | (−0.038, 0.12)  | 0.3084  |
| VSTM: test number 4              | 0.003    | (−0.1, 0.106)   | 0.9529  |
| VSTM: test number 5              | 0.051    | (−0.076, 0.177) | 0.4328  |
| VSTM: test number 6              | 0.009    | (−0.157, 0.175) | 0.9166  |
| VSTM: HR below 120               | −0.009   | (−0.042, 0.025) | 0.6157  |
| VSTM: HR between 120 and 160     | −0.024   | (−0.052, 0.005) | 0.1041  |
| VSTM: HR above 160               | 0.004    | (−0.017, 0.026) | 0.69    |

such as the Grand Canyon provides unique challenges. Participants have little to no interaction with researchers during the approximately 12-hour hike, so devices cannot be easily checked and adjusted. The R2R hike also pushes the battery limits of current commercial off-the-shelf fitness devices, leading to instances of missing or inaccurate data toward the end of the hike. Further fine tuning the devices and simplifying the set-up process will help address some of these concerns. Second, this rich data set allows for the construction of more informative features than the simple ones reported in this initial analysis. For example, we could create more complex predictors from the device data, such as rate of acceleration of heart rate after resting or composite responses from the multiple tests collected.

This initial pilot study allowed us to accomplish two major goals: (1) collect data in an extreme environment from two different populations with a decent sample size and (2) understand the quality and pitfalls of the data collected. The major weaknesses in our current study were lost and/or missing data and the inability to verify the validity of heart rate from participants' data. We will build our understanding of the relationship between physiological, cognitive, and biological data with other predictors of fatigue and performance decrement as we continue to analyze the data collected from the wearable devices, along with the survey and bloodwork data. Another key component

is analyzing the differences between civilians and military warfighters. These data are both supported and funded by government organizations who desire to apply research findings for the benefit of national security.

This article serves as the first report of this research effort and initial analyses. Data will continue to be collected at the Grand Canyon from civilian and military R2R hikers. We anticipate adding and replacing wearable devices and cognitive tasks as well as enhancing our experimental design to increase data quality. We will also explore more rigorous control options (as an example, see the control groups in Kramer et al., 1993). Further data analyses will explore and validate the findings reported in this article and contribute further knowledge to this evolving field of human performance in extreme environments.

### Acknowledgments

This project is funded by the Defense Threat Reduction Agency (DTRA), Project CB10359.

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. SAND2017-12213 J.

## References

- Adolph, E. F., & Associates. (1947). *Physiology of man in the desert*. New York, NY: Interscience.
- Armstrong, L. E., Costill, D. L., & Fink, W. J. (1985). Influence of diuretic-induced dehydration on competitive running performance. *Medicine & Science in Sports & Exercise*, 17(4), 456–461. <https://doi.org/10.1249/00005768-198508000-00009>
- Basner, M., Savitt, A., Moore, T. M., Port, A. M., McGuire, S., Ecker, A. J., ... Gur, R. C. (2015). Development and validation of the cognition test battery for spaceflight. *Aerospace Medicine and Human Performance*, 86(11), 942–952. <https://doi.org/10.3357/AMHP.4343.2015>
- Borreson, J., & Lambert, M. I. (2009). The quantification of training load, the training response, and the effect on performance. *Sports Medicine*, 39(9), 779–795. <https://doi.org/10.2165/11317780-000000000-00000>
- Bourne, L. E., & Yaroush, R. A. (2003). *Stress and cognition: A cognitive psychological perspective* (NAG2-1561). National Aeronautics and Space Administration.
- Buskirk, E. R., Iampietro, P. F., & Bass, D. E. (1958). Work performance after dehydration: Effects of physical conditioning and heat acclimatization. *Journal of Applied Physiology*, 12(2), 189–194. <https://doi.org/10.1152/jappl.1958.12.2.189>
- Cadarette, B. S., Sawka, M. N., Toner, M. M., & Pandolf, K. B. (1984). Aerobic fitness and the hypohydration response to exercise-heat stress. *Aviation, Space, and Environmental Medicine*, 55(6), 507–512.
- Connors, C. K., & Sitarenios, G. (2011). Connors' Continuous Performance Test (CPT). In *Encyclopedia of Clinical Neuropsychology* (pp. 681–683). New York, NY: Springer. [https://doi.org/10.1007/978-0-387-79948-3\\_1535](https://doi.org/10.1007/978-0-387-79948-3_1535)
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87–185. <https://doi.org/10.1017/S0140525X01003922>
- Ellis, H. D. (1982). The effects of cold on the performance of serial choice reaction time and various discrete tasks. *Human Factors*, 24(5), 589–598. <https://doi.org/10.1177/001872088202400509>
- Ellis, H. D., Wilcock, S. E., & Zaman, S. A. (1985). Cold and performance: The effects of information load, analgesics, and the rate of cooling. *Aviation, Space, and Environmental Medicine*, 56(3), 233–237.
- Emmanuel-Aviña, G. E., Abbott, R., Anderson-Bergman, C., Divis, K. M., Jelinkova, L., Newton, V., ... Femling, J. K. (2017). Rim-to-Rim Wearables at the Canyon for Health (R2R WATCH): Experimental design and methodology. *Augmented Cognition, Neurocognition & Machine Learning*, 263–274.
- Enander, A. E. (1989). Effects of thermal stress on human performance. *Scandinavian Journal of Work, Environment, & Health*, 15, 27–33.
- Eriksen, B. A. & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16(1), 143–149. <https://doi.org/10.3758/BF03203267>
- Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., & Posner, M. I. (2005). The activation of attentional networks. *Neuroimage*, 26(2), 471–476. <https://doi.org/10.1016/j.neuroimage.2005.02.004>
- Garigan, T. P., & Ristedt, D. E. (1999). Death from hyponatremia as a result of acute water intoxication in an army basic trainee. *Military Medicine*, 164(3), 234.
- Ghiglieri, M. P., & Myers, T. M. (2001). *Over the edge: Death in the Grand Canyon*. Flagstaff, AZ: Puma Press.
- Grether, W. F. (1973). Human performance at elevated environmental temperatures. *Aerospace Medicine*, 44(7), 747–755.
- Hancock, P. A., & Vasmatazidis, I. (1998). Human occupational and performance limits under stress: The thermal environment as a prototypical example. *Ergonomics*, 41(8), 1169–1191. <https://doi.org/10.1080/001401398186469>
- Hocking, C., Silberstein, R. B., Lau, W. M., Stough, C., & Roberts, W. (2001). Evaluation of cognitive performance in the heat by functional brain imaging and psychometric testing. *Comparative Biochemistry and Physiology: A Molecular and Integrative Physiology*, 128, 719–734. [https://doi.org/10.1016/S1095-6433\(01\)00278-1](https://doi.org/10.1016/S1095-6433(01)00278-1)
- Hyde, D., Thomas, J. R., Schrot, J., & Taylor, W. F. (1997). *Quantification of special operations mission-related performance: Operational evaluation of physical measures*. Bethesda, MD: Naval Medical Research Institute. <https://doi.org/10.21236/ADA381211>
- Leung, D. (2015, November 13). Besides a Stephen Curry advantage, Golden State Warriors flaunt a tech advantage. *The Mercury News*. Retrieved from <http://www.mercurynews.com>.
- Lowe, M., Harris, W., Kane, R. L., Banderet, L., Levinson, D., & Reeves, D. (2007). Neuropsychological assessment in extreme environments. *Archives of Clinical Neuropsychology*, 22S, S89–S99. <https://doi.org/10.1016/j.acn.2006.10.010>
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), 279–281. <https://doi.org/10.1038/36846>
- Karatsoreos, I. N., & McEwen, B. S. (2010). Stress and allostatis. In *Handbook of behavioral medicine: Methods and applications* (pp. 649–658). New York, NY: Springer.
- Kramer, A. F., Coyne, J. T., & Strayer, D. L. (1993). Cognitive function at high altitude. *Human Factors*, 35(2), 329–344.
- Mountain, S. J., & Coyle, E. F. (1992). Influence of graded dehydration on hyperthermia and cardiovascular drift during exercise. *Journal of Applied Physiology*, 73(4), 1340–1350.
- Muza, S. R., Kaminsky, D., Fulco, C. S., Banderet, L. E., & Cymerman, A. (2004). Cysteinyl leukotriene blockade does not prevent acute mountain sickness. *Aviation, Space, and Environmental Medicine*, 75(5), 413–419. <https://doi.org/10.21236/ADA423394>
- Poussel, M., Laroppe, J., Hurdziel, R., Girard, J., Poletti, L., Thil, C., ... Chenuel, B. (2015). Sleep management strategy and performance in an extreme mountain ultra-marathon. *Research in Sports Medicine*, 23(3), 330–336. <https://doi.org/10.1080/15438627.2015.1040916>
- Rastogi, S. K., Gupta, B. N., Husain, T., & Mathur, N. (1990). Cardiac stress in glass bangle workers. *European Journal of Applied Physiology and Occupational Physiology*, 61, 251–257. <https://doi.org/10.1007/BF00357608>
- Rodahl, K. (1989). *The physiology of work*. London, UK: Taylor & Francis.
- Rodahl, K. (2002). Occupational health conditions in extreme environments. *Annals of Occupational Hygiene*, 47(3), 241–252.
- Sawka, M. N., Mountain, S. J., & Latzka, W. A. (2001). Hydration effects on thermoregulation and performance in the heat. *Comparative Biochemistry & Physiology Part A*, 128, 679–690. [https://doi.org/10.1016/S1095-6433\(01\)00274-4](https://doi.org/10.1016/S1095-6433(01)00274-4)
- Sawka, M. N., Toner, M. M., Francesconi, R. P., & Pandolf, K. B. (1983). Hypohydration and exercise: Effects of heat acclimation, gender, and environment. *Journal of Applied Physiology*, 55(4), 1147–1153. <https://doi.org/10.1152/jappl.1983.55.4.1147>
- Sawka, M. N., Young, A. J., Francesconi, R. P., Muza, S. R., & Pandolf, K. B. (1985). Thermoregulatory and blood responses during exercise at graded hypohydration levels. *Journal of Applied Physiology*, 59(5), 1394–1401. <https://doi.org/10.1152/jappl.1985.59.5.1394>
- Strydom, N. B., & Holdsworth, D. L. (1968). The effects of different levels of water deficit on physiological responses during heat stress. *European Journal of Applied Physiology*, 26(2), 95–102. <https://doi.org/10.1007/BF00699507>
- Stück, M., Balzer, H., Hecht, K., & Schröder, H. (2005). Psychological and psychophysiological effects of a high-mountain expedition to Tibet. *Journal of Human Performance in Extreme Environments*, 18(1), 11–20.
- Veltman, J. A., & Gaillard, A. W. K. (1998). Physiological workload reactions to increasing levels of task difficulty. *Ergonomics*, 41(5), 656–669. <https://doi.org/10.1080/001401398186829>
- Wickens, C. D., Keller, J. W., & Shaw, C. (2015). Human factors in high-altitude mountaineering. *Journal of Human Performance in Extreme Environments*, 12(1), 1. <https://doi.org/10.7771/2327-2937.1065>

## Appendix

Not every participant completed the same number of cognitive batteries. Hikers were asked to complete the BrainBaseline tasks at the beginning and end of the hike and then about every three hours in between during natural breaks in the hike. Some participants completed the hike more quickly and thus had fewer opportunities to take the cognitive battery. Some participants waited longer than the recommended three hours to complete the battery. Figure 1 shows the location of the tests (with Test 1 indicating the first time they completed the battery, Test 2 the second, etc.). One benefit of this variance in location and timing of instances of the cognitive battery is that it allows us to tease out some of the learning effects experienced by volunteers. Despite intentionally choosing to include tasks with relatively small learning effects, participants will still tend to get better at the tasks the more times they complete them. Being able to statistically control for this learning effect in our models is helpful. Tables 6–9 show the main effects of interest in addition to breaking down the learning effects. Figures 2–5 graphically represent this information. Notably the confidence intervals tend to increase for later tests due to a smaller sample size (for example, relatively few participants completed 6 sessions of the cognitive battery as opposed to 4 sessions).

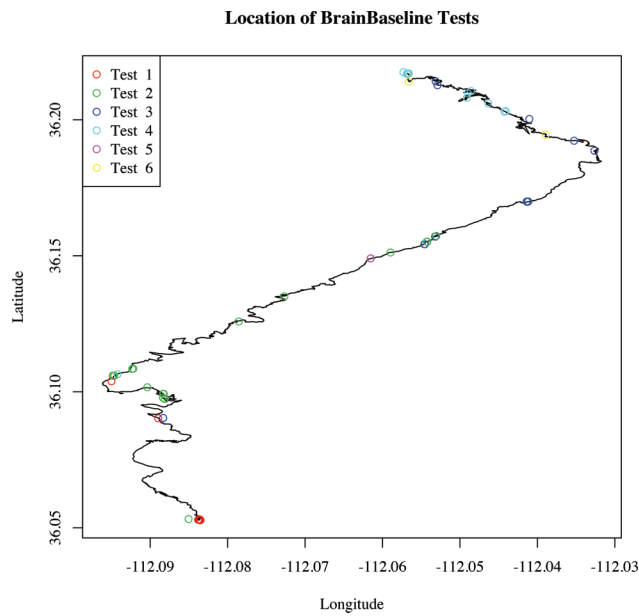


Figure 1. Locations along the R2R trail where BrainBaseline cognitive battery was completed.

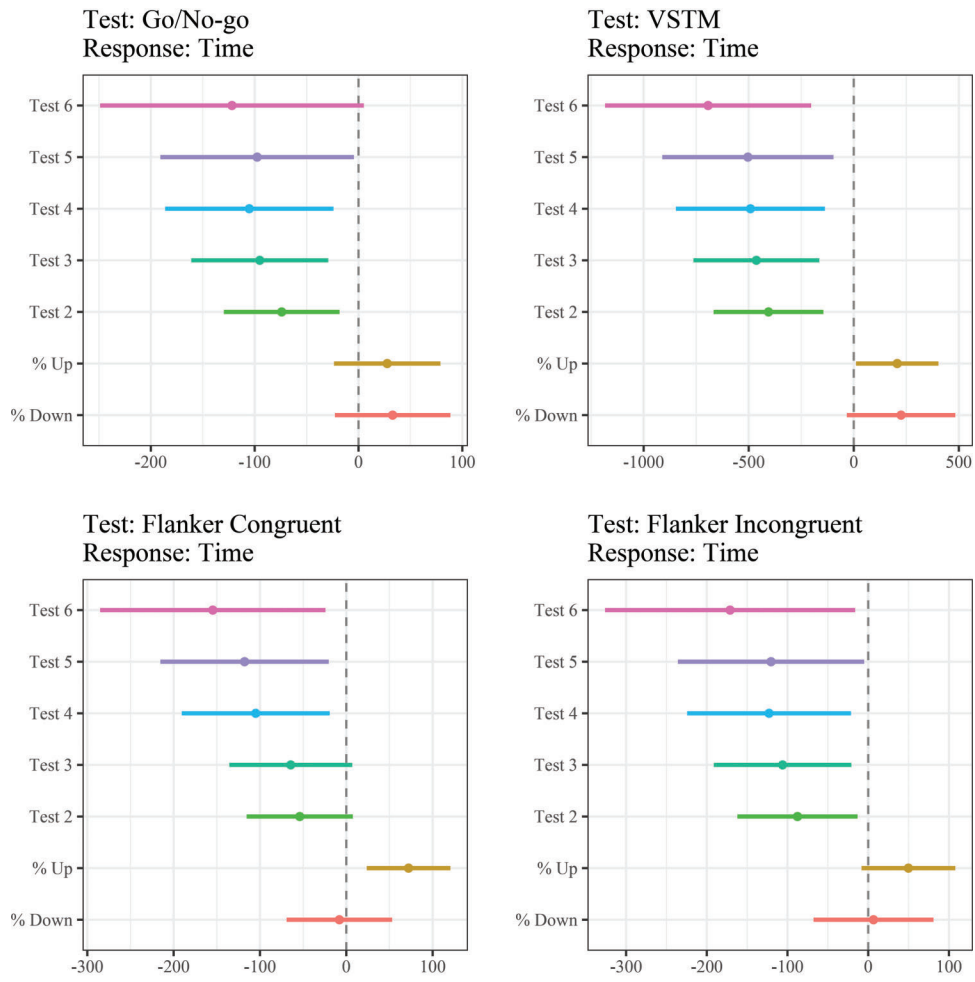


Figure 2. Response time effects for the cognitive battery as a function of proportion up and down the canyon, including learning effects.

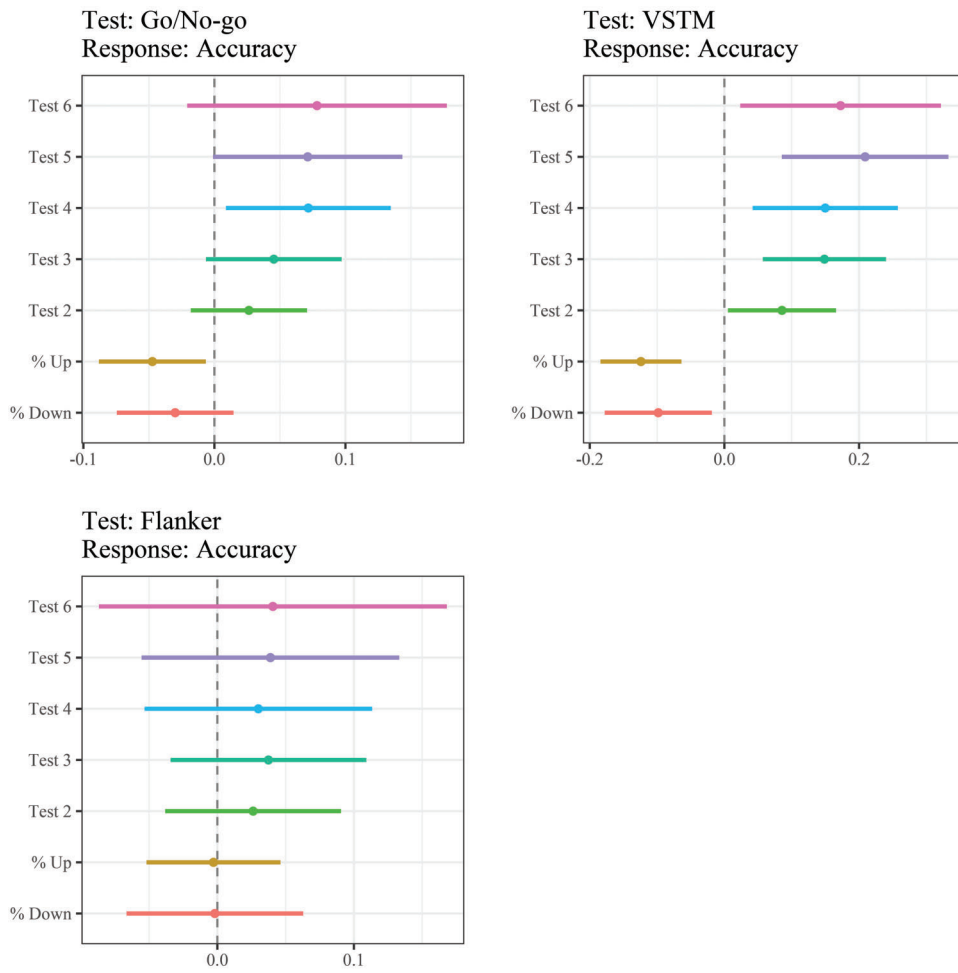


Figure 3. Accuracy effects for the cognitive battery as a function of proportion up and down the canyon, including learning effects.

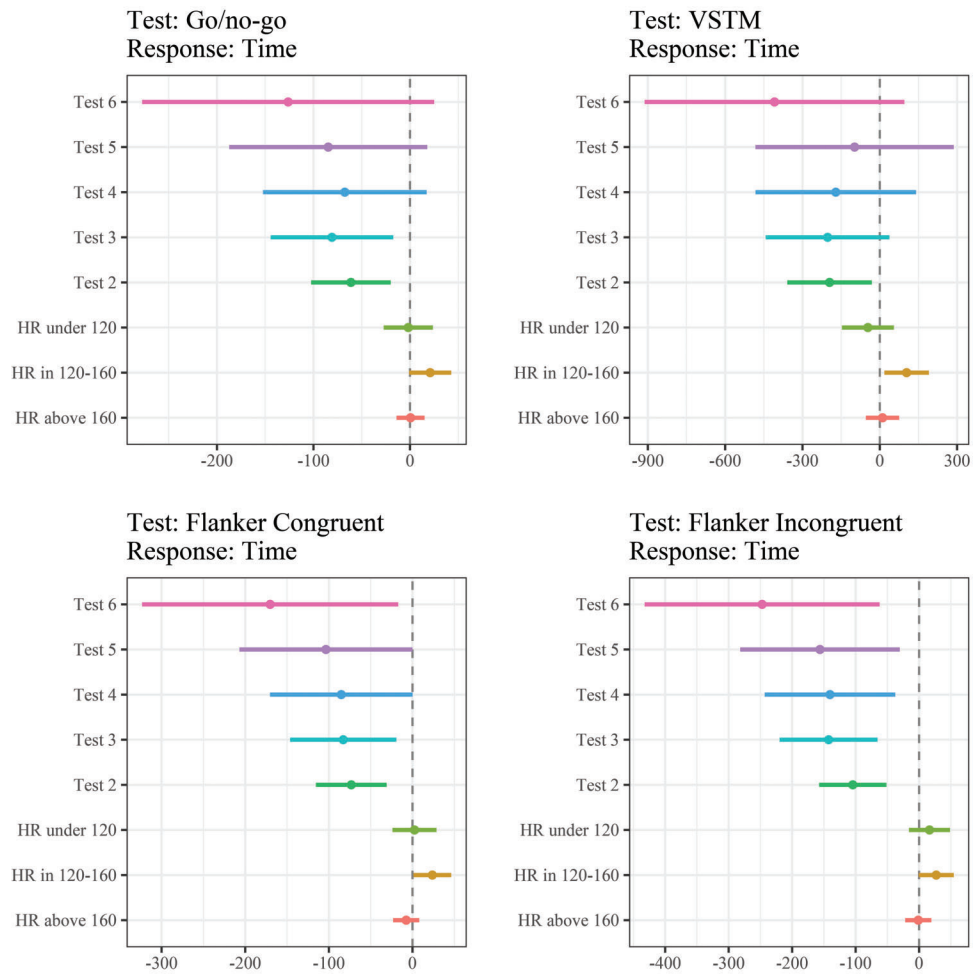


Figure 4. Response time effects for the cognitive battery as a function of heart rate (HR) zone, including learning effects.



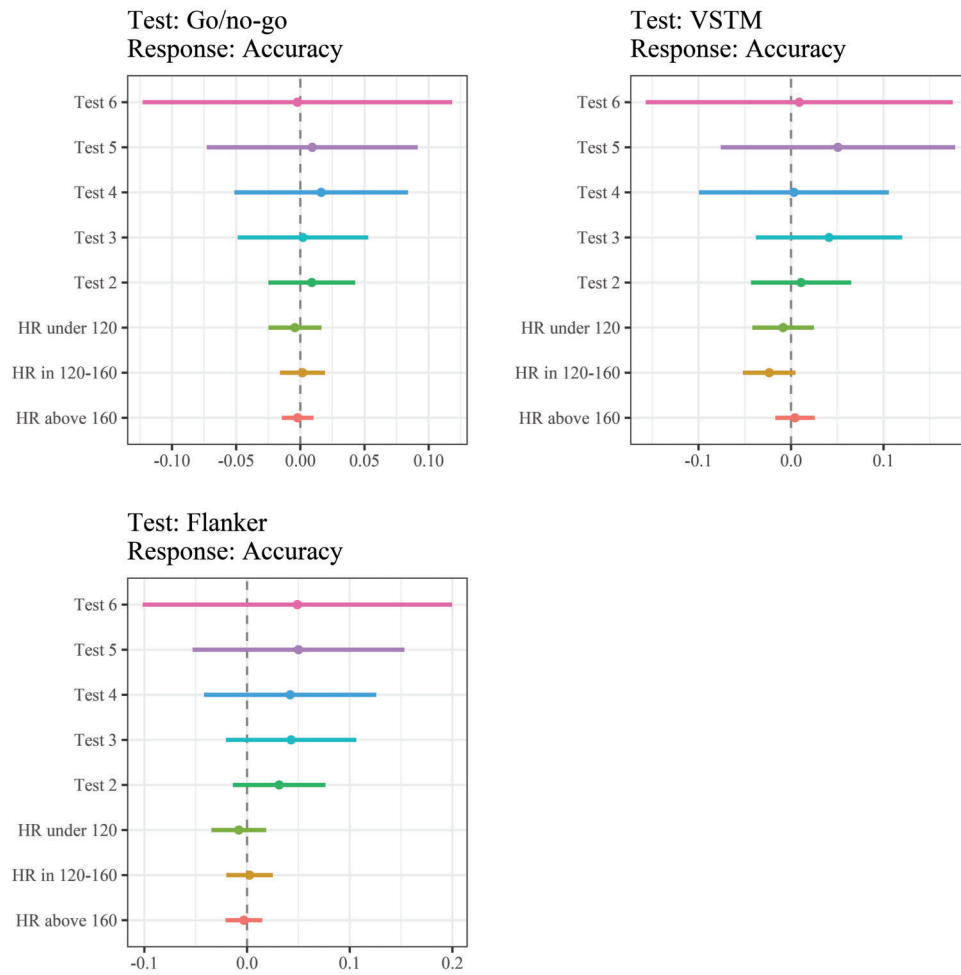


Figure 5. Accuracy effects for the cognitive battery as a function of heart rate (HR) zone, including learning effects.