Physiologic Predictors of Cognitive Performance Variability in Operational Military Environments

From hypothesis to finding: **the variability is large and individually patterned, but no fielded wearable channel predicts it.**

Trevor Barrett, Jill Boberg, Matthew Bartels, Leslie Saxon

Center for Body Computing · USC Institute for Creative Technologies

MHSRS-26-5118 · RESEARCH

USC Institute for Creative Technologies

USC Center for Body Computing

+17.5%
Rise in within-person reaction-time **variability** under operational load, about **4×** the effect on speed

76%
Of service members grew **more erratic** on entering force-on-force operations, beyond any change in speed

5-fold
Spread between steadiest and most erratic service member, a moderately reliable **individual trait**

r = 0.73
The slowest service members are also the **most erratic**: a single shared vulnerability

The Question

Continuous physiological monitoring promises to flag cognitive decline before it degrades mission performance. Most work asks whether service members grow *slower* under load. We asked whether their performance also grows less **reliable** — erratic from one moment to the next — and whether a wearable can anticipate it.

Run-to-run instability may be the more consequential failure mode, yet a mean-based analysis omits it entirely.

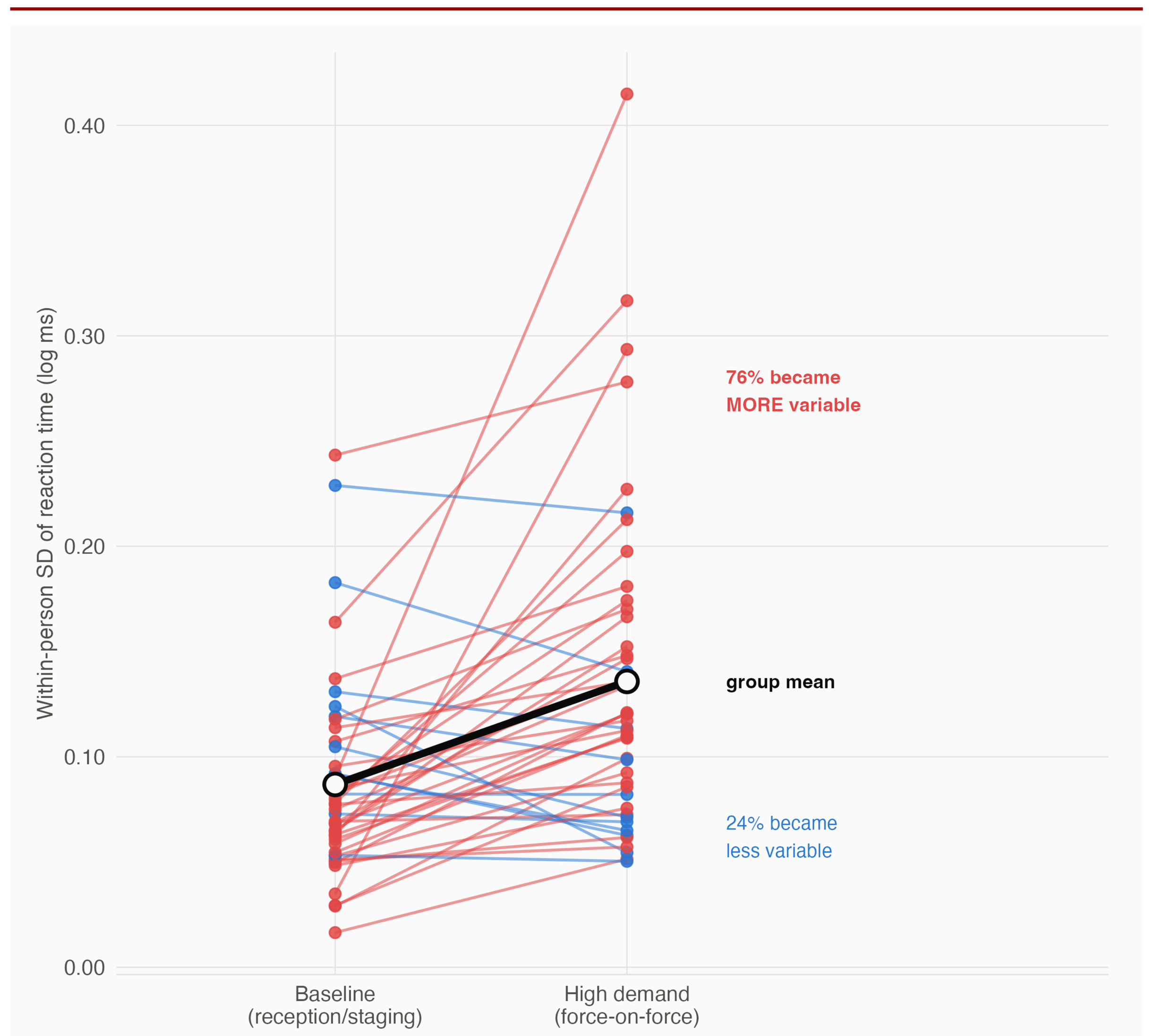
What We Found

Under operational load, service members' reaction times become more variable — a larger change than the shift in average speed — and this vulnerability is a large, stable individual trait. The autonomic sensors we fielded do not detect it.

We hypothesized that nightly autonomic signals, heart-rate variability chief among them, would **predict** these cognitive fluctuations and give early warning of decline. Tested directly, **no autonomic channel (HRV included) predicts the level or the instability of cognition** at this sample size; the effects are equivalence-bounded to zero (Panel 5).

The bounded null is itself informative: the predictive signal is absent, but the phenomenon it was meant to track is real and large. This poster reports both.

The Phenomenon



Each line is one service member's own reaction-time variability, reception/staging versus force-on-force; group mean in black. National Training Center: 45 of 74 service members measured in both phases.

Within-person reaction-time variability rose from **0.087 to 0.139** (log ms), a **60% increase**, moving from reception/staging into force-on-force operations. Paired within-person, **p = 0.0001**; **76% of service members became more variable.**

Design

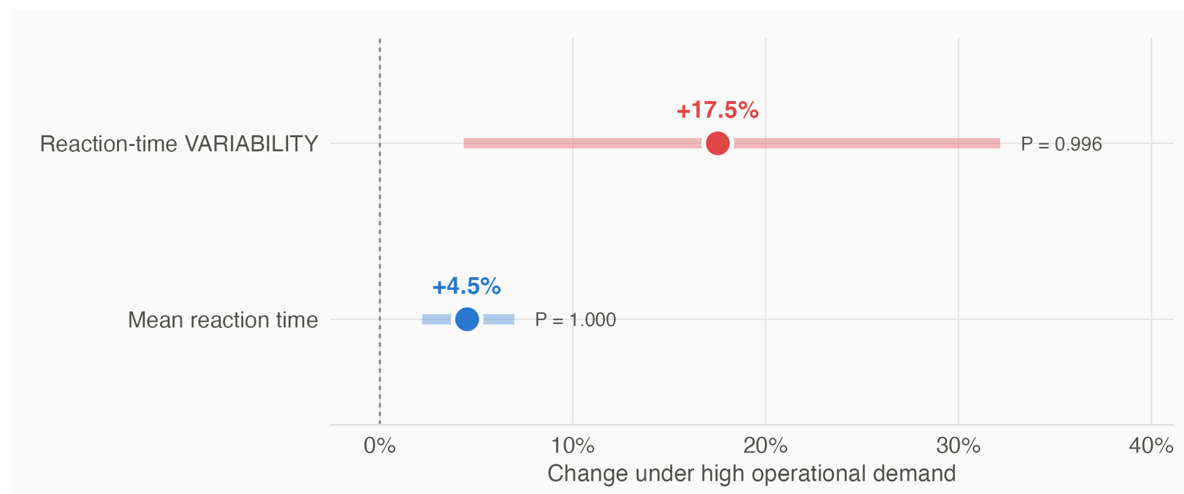
Two analyses, each on the cohort with the repeated measures it requires:

- **Variance arm (National Training Center):** 74 service members, 1,166 psychomotor-vigilance assessments, the densest repeated-measures cognition in the program.
- **Autonomic arm (SOCPAC + Marine close-quarters-battle instructors):** 49 service members, 2,614 assessments, the only cohorts carrying device **heart-rate variability** *and* repeated reaction-time cognition.

Mixed-effects **location-scale** models estimate effects on the mean *and* the within-person variance of reaction time simultaneously. All predictors are within-person; practice-detrended; nulls bounded by equivalence testing.

1 · Variability Rises More Than Speed

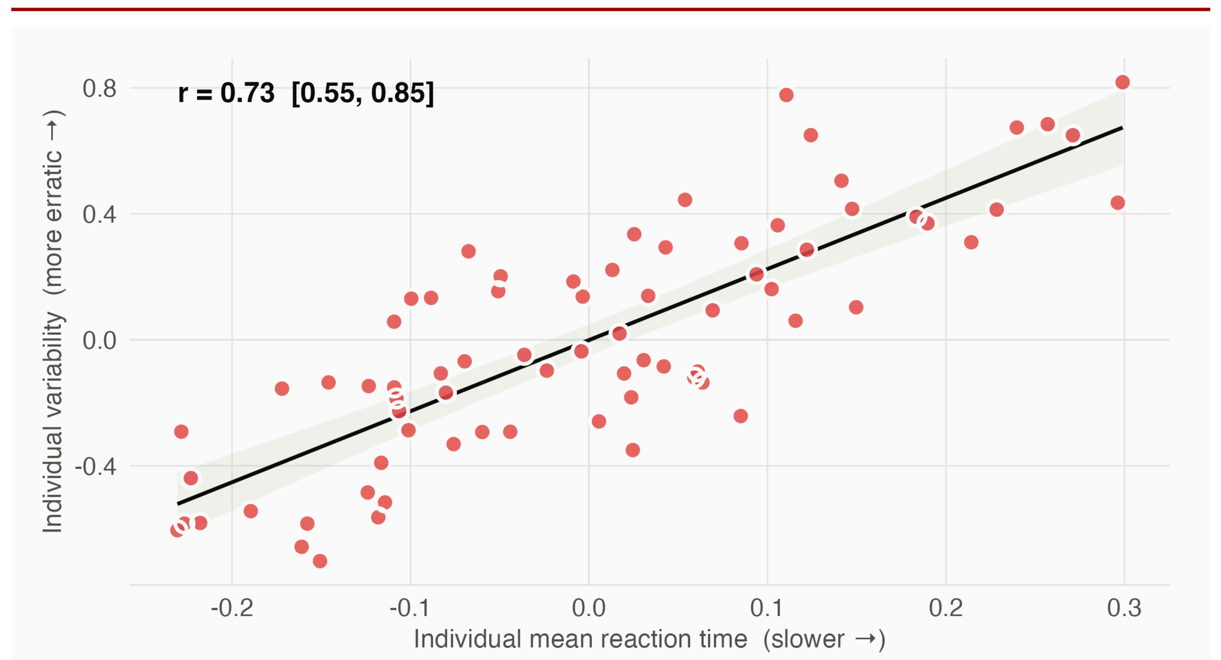
Under load, the mean reaction time rises modestly. The *variability* rises four times as much.



Effect of high demand on the mean reaction time versus its within-person variability, on one axis. Bayesian location-scale model, National Training Center: 74 service members, 1,166 assessments. Bars are 95% credible intervals; P is the posterior probability the effect is positive.

High operational demand raised **mean reaction time by +4.5%** [2.2, 6.9] but **within-person variability by +17.5%** [4.3, 32.2], P = 0.996. **Variability is where the load effect concentrates**, and a mean-only analysis would miss it.

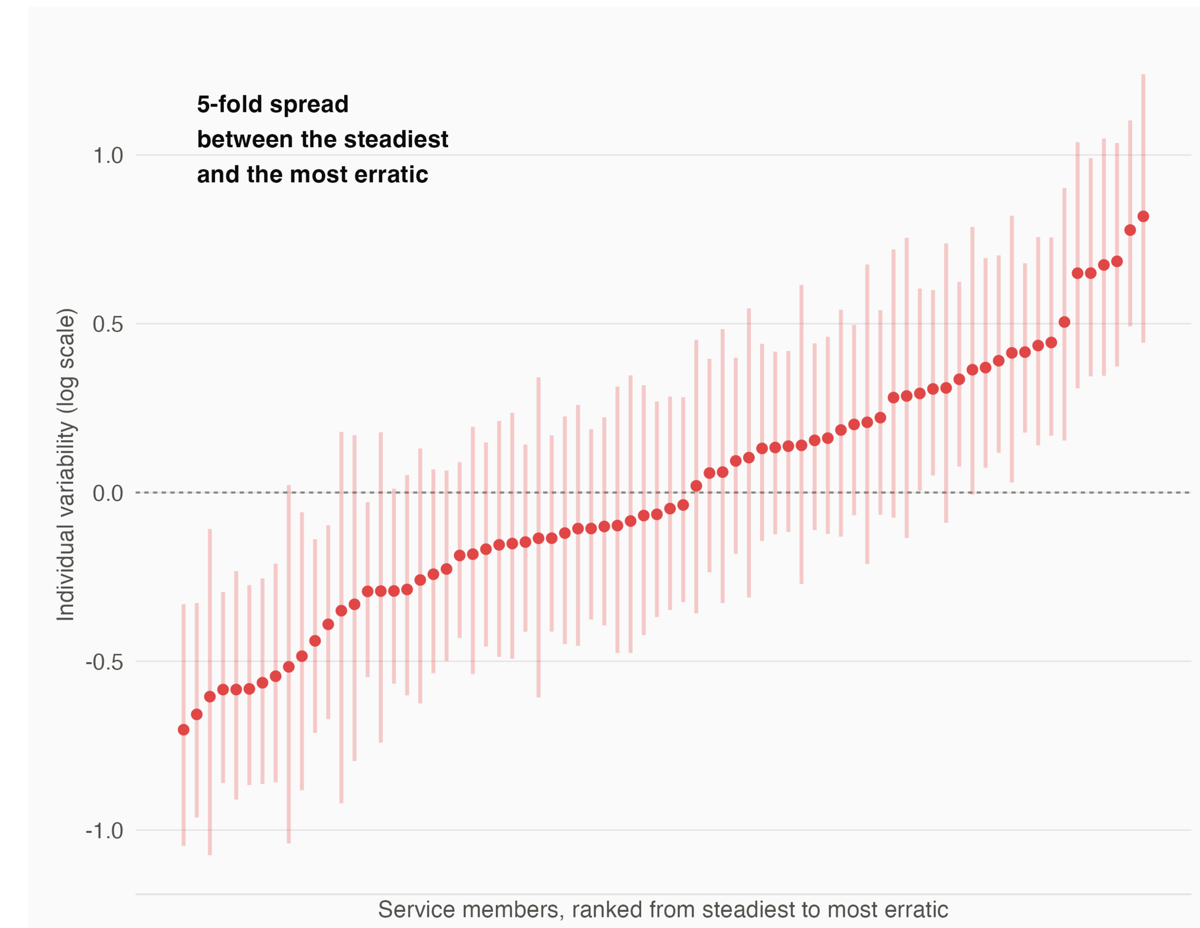
2 · The Slowest Are the Most Variable



Each point is one service member's model random effects, average speed against variability. National Training Center: 74 service members; r is the posterior correlation between the two.

We expected slowness and instability to be independent screening targets. Instead they are strongly correlated, **r = 0.73** [0.55, 0.85]: the **slowest service members** are also the **most variable**. This is consistent with one underlying vulnerability rather than two independent axes.

3 · Who Destabilizes

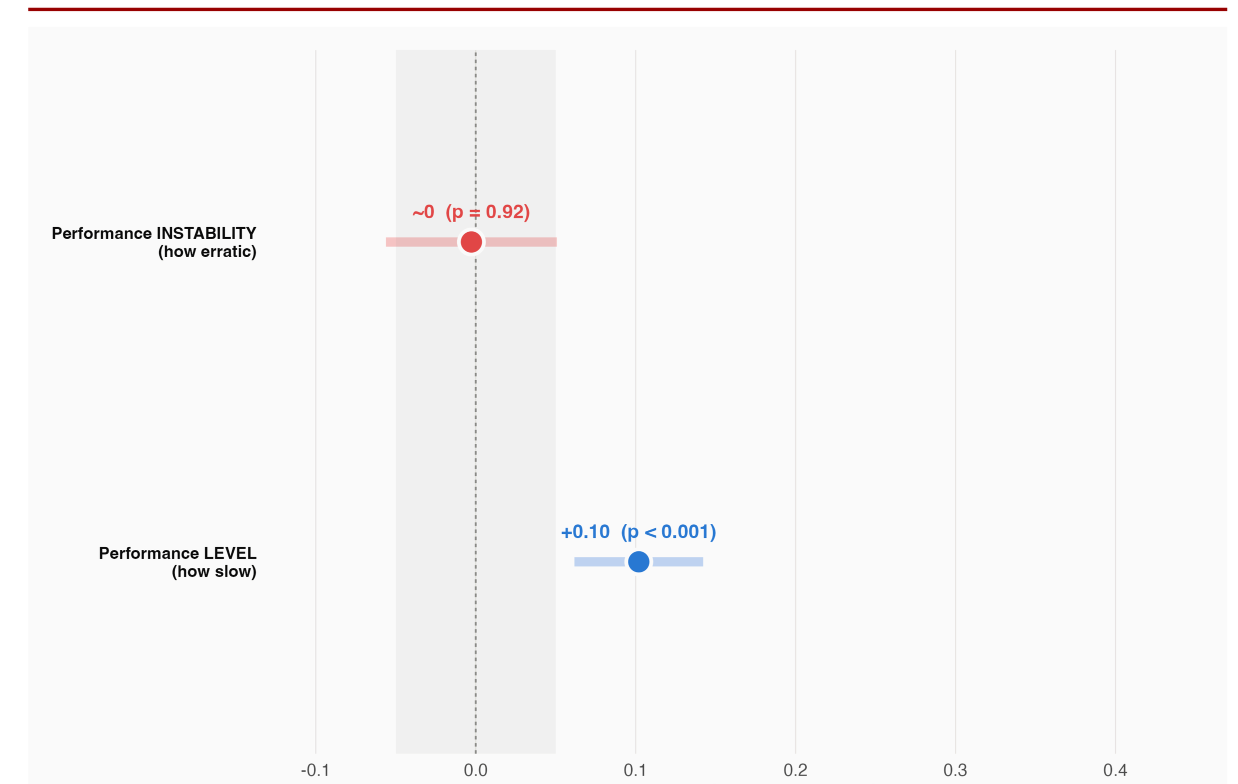


Each service member's own baseline variability, ranked, with 95% intervals. National Training Center: 74 service members.

Service members differ **5-fold** in how erratic they are, and this instability is a **moderately reliable individual trait** (split-half reliability ≈ 0.49; shrinkage-adjusted ≈ 0.81), rather than day-to-day noise. **24% are effectively resilient**, holding steady under a load that destabilizes others.

Implication for monitoring. A population threshold cannot separate the service member who is *always* erratic from the one who is *becoming* erratic, and the second is the case of interest. Referencing each flag to the individual's own baseline is therefore a design requirement.

4 · Strain Predicts Speed, Not Variability

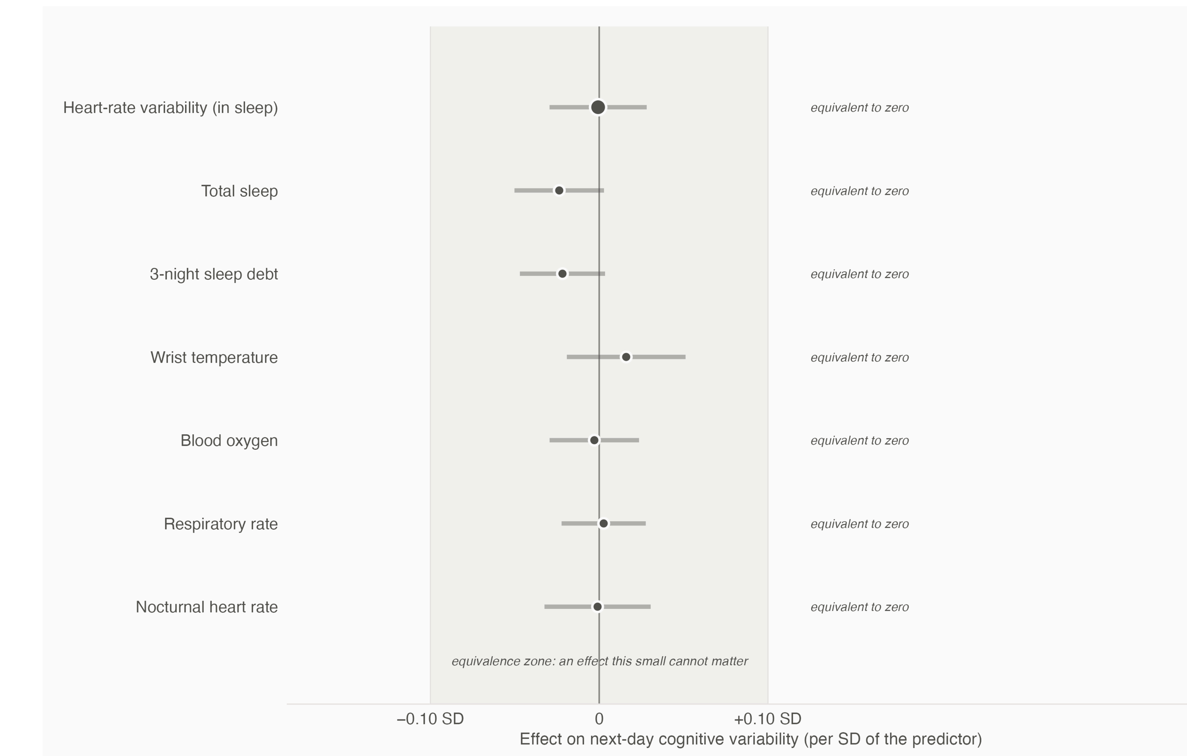


Effect of a measured strain composite (sleep debt, hours awake, sleepiness) on each outcome, in SD units. National Training Center: 74 service members. Instability = reaction-time dispersion after the mean effect is removed.

The measured strain that **strongly predicts a service member's reaction-time level** (+0.10 SD, p < 0.001) has **no relationship to their instability** (≈ 0, p = 0.92). Demand's effect on variability does **not** attenuate when strain is adjusted for (≈ 0%). **Slowness and instability therefore have different drivers.**

5 · No Autonomic Channel Predicts It

No wearable channel, including heart-rate variability (the signal this study was designed around), predicts the instability.



Each channel's effect on next-day cognitive variability, with the ±0.10 SD equivalence zone shaded. SOCPAC + Marine close-quarters-battle instructors: 49 service members, 2,614 assessments; each cognitive test joined to the preceding night's physiology.

Every autonomic and physiologic channel's entire 95% interval falls inside the **±0.10 SD equivalence zone** (≈ ±6 ms of reaction time, below normal trial-to-trial noise): HRV, nocturnal heart rate, respiration, temperature, oxygenation, and sleep. The effects are **formally equivalent to zero** (TOST p < 0.001) — a stronger claim than non-significance. Leave-one-subject-out: **0 of 41** drops significant. The null survived **~18 alternative constructions**: multi-night windows, deviation-from-baseline, autonomic instability, composites, demand moderation, and person-specific coupling.

A limit of the instrument, not a claim about biology. The HRV→cognition effect is real with lab-grade measurement: ECG-HRV tracks PVT decrements under sleep deprivation (Chua et al. 2012), and vagal tone predicts reaction-time *variability* (Williams et al. 2016). Consumer wrist-optical sensors, however, capture the beat-to-beat metrics poorly (Bellenger et al. 2021; Rehman et al. 2024), sampling HRV below the resolution at which those effects replicate: a **capture-resolution** gap.

n = 49 is modestly powered, but the null is *equivalence-bounded*, a stronger statement than "not significant."

What This Means

- Cognitive **instability** under operational load is **large, individually patterned, and distinct from slowing** — a failure mode a mean-based analysis misses.
- Current autonomic and behavioral sensors **do not detect it**. No channel, HRV included, predicts it, and the effects are **equivalence-bounded to zero** rather than simply unproven.
- Detecting *which service members are destabilizing* will require **higher-resolution autonomic acquisition or task-embedded cognitive probes**, not another wrist channel.
- **Knowing what today's wearables cannot measure is itself deployment-relevant:** it informs which capabilities to acquire, and which to avoid.

Acknowledgements. Supported by CDMRP W81XWH-22-1-0956. We thank the service members and the operational units whose participation made this work possible.

References. Chua, *Sleep* 2012 · Williams, *Psychophysiology* 2016 · Bellenger, *Sensors* 2021 · Rehman, *Sensors* 2024