

TECHNICAL ARTICLE 1 OF 2

What workers say is not what the body shows: the bias of self-reported sleepiness in operations

Dersalis R&D team · Dersalis Brasil Inovações em Saúde S.A. · Vitória, Brazil

Over three months of monitoring on the ODN-1 drilling rig, 73 workers completed 2,668 sleepiness self-assessments and 566 objective attention tests. **83% of the self-assessments showed little or no sleepiness. Over the same period, 57% of the attention tests showed low performance.** This gap between what workers perceive and what they can actually deliver is the starting point for measuring fatigue through the body, continuously.

73

workers monitored for 3 months

83%

of self-assessments with no sign of sleepiness

57%

of attention tests with low performance

Why measure fatigue during operations

Human error is cited as the cause of 60% to 90% of accidents and incidents in sectors such as aviation, petrochemicals, steel and transport (Xuecai and Deyong, 2018). Many of these errors arise from an imbalance between a worker's capacity at that moment and the demands of the task. Fatigue is one of the factors that weighs most heavily on this balance: it slows reactions, reduces attention and short-term memory, and impairs judgment (Wong and Swanson, 2022).

The indicators most widely used in industry, however, are reactive: they measure the accident after it has happened (Tang et al., 2018). And human factors assessments tend to rely on the perception of analysts and on the worker's own self-assessment (Ahn et al., 2022). To act before an incident, you need to know how the worker is doing throughout the shift. This study compared three sources of information that could fill that role: self-assessment, an objective attention test and vital signs.

How the study was conducted

Data were collected on the ODN-1 drilling rig, operated by Foresea, during a maintenance stop at a shipyard, as part of an R&D project between Dersalis, Shell and Ocyan. For three months, 73 workers wore a wrist-worn device with an optical sensor and an accelerometer during their shifts, recording heart rate, heart rate variability, skin temperature and movement. A gateway with a cellular connection sent the data to the platform, in compliance with Brazil's data protection law (LGPD).

On the same device, workers answered two types of assessment:

- **Sleepiness self-assessment.** A scale derived from the Karolinska Sleepiness Scale (KSS), the reference for measuring sleepiness subjectively (Martin et al., 2023). To make it easier to use in the field, the scale was simplified from 9 to 5 points, as in the Japanese version validated by Kaida et al. (2006). Point 3 marks the transition between alert and sleepy.
- **Attention test (PVT).** The Psychomotor Vigilance Test measures reaction time to visual stimuli and is considered the behavioral gold standard for detecting loss of vigilance (Khitrov et al., 2014; Kerkamm et al., 2022). The 3-minute version was used, as it fits more easily into the operational routine.

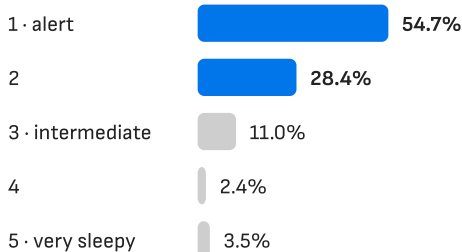
For vital signs, a risk index was adapted from the NEWS 2 protocol (National Early Warning Score), used in hospitals to identify early signs of clinical deterioration (Royal College of Physicians, 2017). Each reading received a score according to how far it was from the range considered normal, and the ranges were color-coded following the Manchester Triage System.

Only data collected during shifts were analyzed, with no intervention from the researchers. Signal segments with motion noise were discarded. In total, the study gathered more than 115,000 heart rate and heart rate variability readings, about 167,000 skin temperature readings, 2,668 self-assessments and 566 attention tests.

What workers say and what the test shows

Sleepiness self-assessment

2,668 answers · scale from 1 (alert) to 5 (very sleepy)



Attention test (3-min PVT)

566 tests · test performance

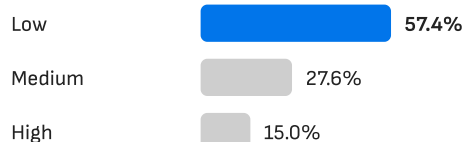


Figure 1. Distribution of sleepiness self-assessments and attention test performance during the study period.

Most workers said they were fine: 83.1% of the answers were at levels 1 and 2 of the scale, with no sign of sleepiness, and only 5.9% at levels 4 and 5. The attention tests told a different story: 57.4% showed low performance, 27.6% medium and only 15.0% high.

The timing did not match either. Self-assessments reporting more sleepiness clustered around 8 p.m. and 6 a.m., close to shift changes. The worst attention test results, on the other hand, clustered between 1 a.m. and 2 a.m.

The most direct reading is that workers tend to underestimate their own fatigue. There is also a second factor. Taken without supervision, in the middle of operations, the attention test can capture more than vigilance: rushing to get back to the task, interruptions and low willingness to take the test also lower the result. In other words, neither measure, on its own, reliably represents the worker's actual state.

What the vital signs showed

Average heart rate during shifts was 82 beats per minute, with most readings between 70 and 90 bpm. Skin temperature averaged 30.2 °C, concentrated between 30 and 33 °C.

When the risk index was applied with clinical reference ranges, designed for the general population, only a small share of the readings fell within the normal range. This does not mean that most workers were at risk. It means that the physical and metabolic load of maintenance work shifts what is normal for that operation. Ranges copied from clinical practice generate too many alerts and lose value as a safety barrier.

The practical conclusion is that alert thresholds need to be adjusted to each operation and, ideally, to each person, taking into account the predominant activity and the physiological limits of the people doing the work.

What this changes in fatigue management

- **Self-assessment alone does not work as a barrier.** It helps engage workers, but it underestimates fatigue precisely when it matters most.
- **The attention test remains a valuable tool.** It is sensitive and objective, and works best when applied with context: at the start of a shift, after an alert or with supervision. Used as a spot check, it helps confirm the worker's state.
- **Continuous physiological data complete the picture.** They do not depend on the worker's answer or interrupt the task, as long as the thresholds are calibrated for the operation.

This combination of findings led Dersalis to develop a fatigue detection model based on heart rate variability, which measures fatigue through the body instead of asking the worker. The method and its field validation are covered in the second article in this series: [Detecting fatigue without interrupting work: the role of parasympathetic activity](#).

Limitations

The study is descriptive and was carried out in a single operation, during a maintenance stop, with tests taken without supervision. Heart rate and temperature ranges were compared with population references, without individual adjustment. The results point to trends that need to be confirmed in other operations and with supervised testing protocols.

References

1. Ahn, S. I., Kurt, R. E. and Akyuz, E. (2022). Application of a SPAR-H based framework to assess human reliability during emergency response drill for man overboard on ships. *Ocean Engineering*, 251, 111089. <https://doi.org/10.1016/j.oceaneng.2022.111089>
2. Kaida, K., Takahashi, M., Åkerstedt, T. et al. (2006). Validation of the Karolinska Sleepiness Scale against performance and EEG variables. *Clinical Neurophysiology*, 117(7), 1574–1581. <https://doi.org/10.1016/j.clinph.2006.03.011>
3. Kerkamm, F., Dengler, D., Eichler, M. et al. (2022). Measurement methods of fatigue, sleepiness, and sleep behaviour aboard ships: a systematic review. *International Journal of Environmental Research and Public Health*, 19(1), 120. <https://doi.org/10.3390/ijerph19010120>

4. Khitrov, M. Y., Laxminarayan, S., Thorsley, D. et al. (2014). PC-PVT: a platform for psychomotor vigilance task testing, analysis, and prediction. *Behavior Research Methods*, 46(1), 140–147. <https://doi.org/10.3758/s13428-013-0339-9>
5. Martin, V. P., Lopez, R., Dauvilliers, Y. et al. (2023). Sleepiness in adults: an umbrella review of a complex construct. *Sleep Medicine Reviews*, 67, 101718. <https://doi.org/10.1016/j.smr.2022.101718>
6. Royal College of Physicians (2017). *National Early Warning Score (NEWS) 2: standardising the assessment of acute-illness severity in the NHS*. London: RCP.
7. Tang, K. H. D., Dawal, S. Z. M. and Olugu, E. U. (2018). A review of the offshore oil and gas safety indices. *Safety Science*, 109, 344–352. <https://doi.org/10.1016/j.ssci.2018.06.018>
8. Wong, I. and Swanson, N. (2022). Approaches to managing work-related fatigue to meet the needs of American workers and employers. *American Journal of Industrial Medicine*, 65(11), 827–831. <https://doi.org/10.1002/ajim.23402>
9. Xuecai, X. and Deyong, G. (2018). Human factors risk assessment and management: process safety in engineering. *Process Safety and Environmental Protection*, 113, 467–482. <https://doi.org/10.1016/j.psep.2017.11.018>

Revised version of Technical Paper 1 (2024), prepared by Dersalis in an R&D project with Shell Brasil and Ocyan. Worker data was handled anonymously, in compliance with Brazil's data protection law (LGPD).