Impact of Sleep Deprivation on Psychological Metrics

Among amateur-friendly metrics, reaction time lapses measured by PVT show the strongest degradation (3-8x increase) during sleep deprivation, followed by declining DSST performance.

Abstract

Psychomotor Vigilance Test (PVT) metrics consistently demonstrate marked degradation in response to sleep deprivation. Reaction time lapses—defined as responses exceeding 500 milliseconds—increase three- to eightfold after 16–22 hours of total sleep deprivation. In addition, inverse reaction time measures decline appreciably within 22–28 hours, though their calculation is less straightforward. The Digit Symbol Substitution Test (DSST) also shows declines in correct responses that correlate with sleep loss and can be measured rapidly with minimal equipment. Other measures, such as log-transformed signal-to-noise ratio and driving performance, register large effects but require specialized methods. Thus, among easily administered tests suitable for nonexperts, simple counts of PVT lapses and DSST correct responses provide sensitive indicators of the effects of sleep deprivation.

Paper search

Using your research question "Which psychological metrics (especially easy and fast to measure ones for amateurs, like digit span or reaction speed) degrade most strongly with sleep deprivation?", we searched across over 126 million academic papers from the Semantic Scholar corpus. We retrieved the 50 papers most relevant to the query.

Screening

We screened in papers that met these criteria:

- **Outcome Measures**: Does the study measure cognitive or psychological performance using quantitative measurements both before and during/after sleep deprivation?
- **Sleep Deprivation Protocol**: Does the study implement a sleep deprivation protocol where participants get less than 4 hours of sleep?
- **Participant Characteristics**: Does the study include only healthy adults (18+ years) without any pre-existing sleep disorders, psychiatric conditions, or neurological conditions?
- **Assessment Methods**: Does the study include at least one psychological metric that can be administered without professional expertise?
- **Baseline Measurements**: Does the study include baseline (well-rested) measurements for comparison?
- **Study Focus**: Does the study examine psychological/cognitive outcomes (rather than exclusively physical performance metrics)?
- **Study Type**: Is the study either a primary research study OR a systematic review/meta-analysis meeting all other criteria?

We considered all screening questions together and made a holistic judgement about whether to screen in each paper.

Data extraction

We asked a large language model to extract each data column below from each paper. We gave the model the extraction instructions shown below for each column.

- **Study Design**:

  Describe the overall experimental design. Specifically note:

- Type of study design (e.g., repeated measures, cross-sectional)
- Whether it involved total sleep deprivation or partial sleep restriction
- Specific experimental conditions (e.g., different sleep duration groups)

Look in the methods or design section. If multiple design elements exist, list all. If design is not clearly stated, write "Not clearly reported".

- **Participant Characteristics**:

  Extract the following participant details:

- Total number of participants
- Gender breakdown
- Age range
- Inclusion/exclusion criteria

Prioritize reporting exact numbers. If ranges are given, report both minimum and maximum. If any characteristic is not reported, write "Not reported".

- **Sleep Restriction/Deprivation Protocol**:

  Describe the specific sleep manipulation:

- Type of sleep manipulation (total or partial sleep deprivation)
- Duration of sleep restriction/deprivation
- Specific sleep conditions (e.g., hours in bed per night)
- Total duration of experimental period

Be precise about hours and days. If multiple conditions exist, list all. Use exact numbers from the text.

- **Psychological Metrics Measured**:

  List all psychological performance metrics used:

- Specific tests/tasks employed
- Metrics within those tests (e.g., reaction time, lapses)
- Measurement frequency
- Duration of each measurement

Focus on metrics related to cognitive performance, especially those easily measured by amateurs. Prioritize metrics directly related to reaction time, vigilance, and cognitive speed.

- **Key Performance Outcomes**:

  Extract the primary findings for each psychological metric:

- Specific changes observed during sleep deprivation

- Magnitude of performance degradation
- Any recovery or adaptation patterns

Look in results section. Report numerical changes if available. If only qualitative descriptions exist, quote directly from text.

- **Study Setting**:

  Describe the research environment:

- Laboratory type (e.g., controlled laboratory)
- Geographic location
- Specific environmental controls

If minimal details are provided, write "Controlled laboratory setting" or "Not specified".

Results

Characteristics of Included Studies

Study	Study Design	Metrics Measured	Sleep Deprivation Protocol	Sample Size	Full text retrieved
Basner and Dinges, 2011	Repeated measures	Psychomotor Vigilance Test (PVT) (lapses, mean 1/reaction time (RT), mean slowest 10% 1/RT)	Total Sleep Deprivation (TSD): 33h awake; Partial Sleep Deprivation (PSD): 5 nights of 4h Time in Bed (TIB)	74	No
Belenky et al., 2003	Repeated measures	PVT (speed, lapses)	PSD: 7 days of 3h, 5h, 7h, or 9h TIB	66	No
Brieva et al., 2021	Repeated measures	Digit Symbol Substitution Test (DSST), PVT10 (lapses, 1/RT), Karolinska Sleepiness Scale (KSS)	PSD: 5 nights of 4h TIB; TSD: 36h	41	No

Study	Study Design	Metrics Measured	Sleep Deprivation Protocol	Sample Size	Full text retrieved
Chavali et al., 2017	Repeated measures	PVT (Signal-to-Noise Ratio (SNR), Log-transformed SNR (LSNR), LSNR approximation (LSNRapx), lapses)	TSD: 38h	99	Yes
Drummond et al., 2005	Repeated measures	PVT (RT, lapses), Visual Analogue Scale	TSD: 36h	20	Yes
Jongen et al., 2015	Crossover	PVT, Divided Attention Test (DAT), Attention Network Test (ANT), DSST, Useful Field of View (UFOV)	TSD: 26h	24	Yes
Patanaik et al., 2015	Repeated measures	PVT (RT, lapses, response speed)	TSD: Dataset 1: ~9.5h; Dataset 2: 26h	180	Yes
Rakitin et al., 2012	Repeated measures	PVT (RT, lapses), Delayed Letter Recognition (DLR) task	TSD: 48h	26 (20 for DLR, 21 for PVT)	Yes
Scott et al., 2006	Crossover	Simple and choice RT, tracking task, number cancellation, Profile of Mood States (POMS)	TSD: 30h	6	No
Tucker et al., 2009	Repeated measures	PVT (mean RT, lapses, false starts)	TSD: 54h	84	No

Based on our analysis of the 10 included studies:

- Study Design : 8 studies reported using a repeated measures design, while 2 reported using a crossover

design.

- Metrics Measured :
  - The Psychomotor Vigilance Test (PVT) was reported as a metric in 9 out of the 10 studies.
  - The Digit Symbol Substitution Test (DSST) was used in 2 studies.
  - Various other tests (e.g., KSS, Visual Analogue Scale, DAT, ANT, UFOV, DLR, RT, tracking task, number cancellation, and POMS) were each reported in 1 study.
- Sleep Deprivation Protocol :
  - 7 studies reported using only total sleep deprivation (TSD)
  - 1 study reported using only partial sleep deprivation (PSD)
  - 2 studies reported incorporating both TSD and PSD
  - The duration of TSD ranged from 26 to 54 hours
  - PSD protocols varied in duration and sleep restriction
- Sample Size :
  - Sample sizes ranged from 6 to 180 participants
  - The total number of participants across all studies was 620
  - There was considerable variation in sample sizes, with some studies having fewer than 30 participants and others having over 100

Effects of Sleep Deprivation on Psychological Metrics

Primary Performance Metrics

Metric	Description	Sensitivity to Sleep Deprivation	Ease of Measurement
PVT Lapses	Reaction times $>$ 500ms	High; consistent increases across studies	High; simple count
PVT Mean 1/RT	Inverse of mean reaction time	High; consistent decreases	Moderate; requires calculation
PVT Mean Slowest 10% 1/RT	Inverse of mean of slowest 10% RTs	High; large effect sizes	Moderate; requires data processing
DSST Correct Responses	Number of correct symbol substitutions	Moderate to High; consistent decreases	High; simple count
Standard Deviation of Lateral Position (SDLP) (Driving Performance)	Standard Deviation of Lateral Position	High; significant increase (3.1 cm)	Low; requires specialized equipment
LSNR	Log-transformed Signal-to-Noise Ratio	High; 36-50% reduction in fidelity	Low; requires complex calculation

Our analysis of the 6 metrics for their sensitivity to sleep deprivation and ease of measurement suggests:

- Sensitivity to Sleep Deprivation :
  - High sensitivity for 5 out of 6 metrics

  - Moderate to high sensitivity for 1 out of 6 metrics
- Ease of Measurement :
  - High ease of measurement for 2 out of 6 metrics
  - Moderate ease of measurement for 2 out of 6 metrics
  - Low ease of measurement for 2 out of 6 metrics
- Metrics with both high sensitivity and high ease of measurement :
  - PVT Lapses (reaction times > 500ms)
  - DSST Correct Responses (number of correct symbol substitutions)
- PVT-based metrics :
  - All PVT-based metrics (Lapses, Mean 1/RT, Mean Slowest 10% 1/RT) showed high sensitivity to sleep deprivation
  - Only PVT Lapses had high ease of measurement, while the others were moderate
- Other metrics :
  - SDLP (Driving Performance) and LSNR (Log-transformed Signal-to-Noise Ratio) both showed high sensitivity but low ease of measurement, likely due to the specialized equipment or complex calculations required

Comparative Degradation Rates

Metric Type	Degradation Rate	Time to Significant Effect	Measurement Complexity
PVT Lapses	High; 3-8x increase in vulnerable individuals	Detectable within 16-22 hours of TSD	Low
PVT Mean 1/RT	Moderate to High; effect sizes 0.88-1.94	Detectable within 22-28 hours of TSD	Low to Moderate
DSST Correct Responses	Moderate; significant correlations with sleep loss	Detectable within 22-28 hours of TSD	Low
LSNR	High; 36-50% reduction in 22-38 hours	Detectable within 22 hours of TSD	High
Driving Performance (SDLP)	Moderate; 3.1 cm increase after TSD	Detectable after 26 hours of TSD	High

Our analysis of 5 different metrics for measuring the effects of sleep deprivation suggests:

- Degradation Rate :
  - High degradation rates for 2 out of 5 metrics (PVT Lapses and LSNR)
  - Moderate to high degradation for 1 out of 5 metrics (PVT Mean 1/RT)
  - Moderate degradation for 2 out of 5 metrics (DSST Correct Responses and Driving Performance)
- Time to Significant Effect :

  - Earliest detectable effect at 16-22 hours of total sleep deprivation (TSD) for PVT Lapses
  - 3 out of 5 metrics showed effects within 22-28 hours of TSD
  - Latest detectable effect at 26 hours of TSD for Driving Performance
- Measurement Complexity :
  - Low complexity for 2 out of 5 metrics (PVT Lapses and DSST Correct Responses)
  - Low to moderate complexity for 1 out of 5 metrics (PVT Mean 1/RT)
  - High complexity for 2 out of 5 metrics (LSNR and Driving Performance)
- The PVT Lapses metric appears to be the most sensitive, showing high degradation and the earliest detection time, while also having low measurement complexity.

Implementation Considerations

Measurement Reliability

Based on our analysis, the reliability of measurements across studies appeared to be generally high, particularly for well-established tasks like the PVT. The use of repeated measures designs in most studies helped to control for individual variability and increase reliability. However, some factors that may affect measurement reliability include:

1. Time-on-task effects : Tucker et al. (2009) reported a significant interaction between sleep deprivation and time-on-task for mean reaction time, suggesting that test duration may influence results.
2. Learning effects : Brieva et al. (2021) noted potential learning effects in the DSST, which could complicate the interpretation of results, especially in repeated testing scenarios.
3. Individual differences : Patanaik et al. (2015) identified significant individual differences in vulnerability to sleep deprivation effects, which may affect the reliability of group-level analyses.
4. Circadian effects : Several studies noted the influence of circadian rhythms on performance, which should be considered when interpreting results and planning measurement timing.

Practical Application

Metric	Required Equipment	Administration Time	Technical Expertise Needed
Psychomotor Vigilance Test (PVT)	Computer/smartphone with millisecond timing	5-10 minutes	Low
Digit Symbol Substitution Test (DSST)	Paper/pencil or computer	2-3 minutes	Low
Simple Reaction Time (RT)	Computer/smartphone with millisecond timing	2-5 minutes	Low
Standard Deviation of Lateral Position (SDLP)	Driving simulator or instrumented vehicle	30-60 minutes	High

Log-transformed Signal-to-Noise Ratio (LSNR)	Computer with specialized software	10 minutes (PVT-based)	High

Our analysis of 5 different metrics for assessing fatigue or alertness suggests:

- Equipment :
  - 3 out of 5 metrics required a computer or smartphone
  - 1 out of 5 could use either paper/pencil or a computer
  - 1 out of 5 required specialized equipment (driving simulator or instrumented vehicle)
  - 1 out of 5 required a computer with specialized software
- Administration Time :
  - 3 out of 5 metrics took 10 minutes or less to administer
  - 1 out of 5 took between 2-5 minutes
  - 1 out of 5 required a longer time, between 30-60 minutes
- Technical Expertise :
  - 3 out of 5 metrics required low technical expertise
  - 2 out of 5 required high technical expertise
- The Psychomotor Vigilance Test (PVT), Digit Symbol Substitution Test (DSST), and Simple Reaction Time (Simple RT) all required low technical expertise and could be administered relatively quickly (10 minutes or less).
- The Standard Deviation of Lateral Position (SDLP) required the most specialized equipment and longest administration time.
- The Log-transformed Signal-to-Noise Ratio (LSNR) required specialized software and high technical expertise but could be administered in a relatively short time.

References

A. Patanaik, C. Kwoh, E. C. Chua, J. Gooley, and M. Chee. "Classifying Vulnerability to Sleep Deprivation Using Baseline Measures of Psychomotor Vigilance." *Sleep*, 2015.

Adrienne M. Tucker, R. Basner, Y. Stern, and B. Rakitin. "The Variable Response-Stimulus Interval Effect and Sleep Deprivation: An Unexplored Aspect of Psychomotor Vigilance Task Performance." *Sleep*, 2009.

B. Rakitin, Adrienne M. Tucker, R. Basner, and Y. Stern. "The Effects of Stimulus Degradation After 48 Hours of Total Sleep Deprivation." *Sleep*, 2012.

G. Belenky, N. Wesensten, D. Thorne, Maria L. Thomas, H. Sing, D. Redmond, M. Russo, and T. Balkin. "Patterns of Performance Degradation and Restoration During Sleep Restriction and Subsequent Recovery: A Sleep Dose-response Study." *Journal of Sleep Research*, 2003.

J. Scott, L. Mcnaughton, and R. Polman. "Effects of Sleep Deprivation and Exercise on Cognitive, Motor Performance and Mood." *Physiology and Behavior*, 2006.

M. Basner, and D. Dinges. "Maximizing Sensitivity of the Psychomotor Vigilance Test (PVT) to Sleep Loss." *Sleep*, 2011.

S. Drummond, A. Bischoff-Grethe, D. Dinges, L. Ayalon, S. Mednick, and M. J. Meloy. "The Neural Basis of the Psychomotor Vigilance Task." *Sleep*, 2005.

S. Jongen, J. Perrier, E. Vuurman, J. Ramaekers, and A. Vermeeren. "Sensitivity and Validity of Psychometric Tests for Assessing Driving Impairment: Effects of Sleep Deprivation." *PLoS ONE*, 2015.

Tess E. Brieva, Caroline A. Antler, Erika Yamazaki, Courtney E Casale, and N. Goel. "127 Cognitive Throughput, Behavioral Attention, and Sleepiness Show Robust Relationships During Sleep Loss but Not During Recovery," 2021.

V. P. Chavali, S. Riedy, and H. V. Van Dongen. "Signal-to-Noise Ratio in PVT Performance as a Cognitive Measure of the Effect of Sleep Deprivation on the Fidelity of Information Processing." *Sleep*, 2017.