Human Interaction and Emerging Technologies (IHIET-AI), Vol. 201, 2026, 226–237
https://doi.org/10.54941/ahfe1007176

When Time Disappears: Uncovering Stress in an Analog Underground Mission

Raphaëlle Giguère[1,2], Caroline Rhéaume[1,2], Christian Clot[3], Marion Trousselard[3], Victor Niaussat[4], Marie-Pierre Gagnon[1,2], Maxime Sasseville[1,2], and Alexandre Marois[1,2]

[1]Université Laval, Québec, QC, G1V 0A6, Canada
[2]VITAM, Québec, QC, G1J 2G1, Canada
[3]Human Adaptation Institute, Marseille, 13002, France
[4]Thales AVS, Mérignac, 33700, France

ABSTRACT

Stress is a central component of human adaptation, particularly in isolated confined extreme (ICE) environments. In such settings, stressors may impair cognitive performance, emotional regulation, decision-making processes, and overall psychological and physical well-being. ICE environments also provide unique opportunities to investigate the boundaries of human adaptability, particularly when they involve temporal and social isolation. This study examined whether heart rate variability (HRV)-derived features could predict perceived stress in a controlled laboratory setting, as well as whether these models could be applied to an analogue ICE experiment without access to self-reported stress assessment. Supervised classification models were trained on the SPACE dataset using HRV features and ratings from the Visual Analogue Scale for Stress (VAS-S). The best-performing models were then applied, without retraining, to physiological data collected during the 15-day DeepTime II cave isolation mission. In the absence of subjective labels, the validity was examined using Baevsky's Stress Index (BSI) as an autonomic reference marker. There was substantial variation in HRV-based models between individuals in the SPACE dataset, and models performed only marginally better than chance at differentiating stress from no-stress conditions. Despite substantial class overlap, predicted stress proxies exhibited descriptive differences in BSI across predicted categories, with higher predicted classes tending to show higher autonomic strain. In the absence of subjective assessments, cardiac autonomic indicators alone provide limited inference of perceived stress, particularly when models are applied to fundamentally different contexts. These findings highlight the constraints of generalized HRV-based stress modelling and support the need for individualized and multimodal approaches in ICE environments.

Keywords: Isolated confined extreme (ICE), Heart rate variability, Perceived stress, Supervised machine learning, Model transferability

INTRODUCTION

Stress is a universal experience that accompanies the challenges of everyday life and represents one of the major risk factors for mental illness (Vaessen *et al.*, 2021). Bernard highlighted the critical role of homeostasis in ensuring organism survival amid environmental changes (Bernard, 1957). Building on

Received February 19, 2026; Revised March 12, 2026; Accepted March 25, 2026; Available online April 1, 2026

this concept, Selye introduced the concept of “stress” to denote factors that could disturb physiological balance (Selye, 1956). While the stress response has evolved to promote adaptation, prolonged or intense stress responses can result in tissue damage and chronic disease (Selye, 1978).

The stress response involves two main physiological pathways. The first is the hypothalamic-pituitary-adrenal (HPA) axis, which regulates cortisol release. The second is the sympathetic branch of the autonomic nervous system (ANS), responsible for the “fight-or-flight” response by releasing noradrenaline (Curtis *et al.*, 1997; Marques *et al.*, 2010). These two pathways interact closely and dynamically (Rotenberg and McGrath, 2016). In response to a stressor, the ANS induces measurable physiological changes, including alterations in heart rate variability (HRV; Kim *et al.*, 2018). HRV reflects the heart’s responsiveness to various physiological and environmental stimuli (Rajendra Acharya *et al.*, 2006). HRV is a good index of ANS function for evaluating a stress response because of its sensitivity to the sympathetic and parasympathetic nervous system (Kim *et al.*, 2018).

As a result, stress is commonly assessed through physiological indicators such as HRV. Baevsky (Baevsky and Chernikova, 2017) developed the Baevsky’s stress index (BSI), an HRV-based marker reflecting sympathetic predominance and autonomic activity. However, there is currently no universally accepted gold standard for stress evaluation (Kim *et al.*, 2018). HRV alone may not accurately capture perceived or psychological stress (Martinez *et al.*, 2022). Subjective assessments combined with physiological markers are therefore considered a promising approach to capture both perceived stress and autonomic activity (Vaessen *et al.*, 2021). Validated tools such as the Visual Analogue Scale for Stress (VAS-S; Lesage *et al.*, 2012) provide a simple and rapid self-reported measure of perceived stress.

However, subjective stress assessments may be impractical or impossible in extreme and naturalistic settings. Rather than representing a methodological weakness, the absence of self-report can be viewed as a structural constraint inherent to certain environments, highlighting the need for methods capable of inferring perceived stress from physiological signals. This challenge is particularly relevant in Isolated, Confined, and Extreme (ICE) environments, characterized by isolation, limited space, discomfort, and reduced privacy (Manzey *et al.*, 1998). Caves exemplify a typical ICE environment, exposing individuals to stressors, including confinement, physical discomfort, disrupted photoperiods (i.e. light-dark cycle), lack of privacy, altered temporal perception, limited communication, sensory deprivation, monotony, and high humidity (Le Roy *et al.*, 2023). Similar prolonged cave isolation protocols have been used to investigate physiological adaptation to extreme subterranean confinement. For example, extended cavern isolation has been shown to induce measurable physiological changes supporting deep caves as valid analogue environments for studying human adaptation in the absence of natural temporal cues (Arbeille *et al.*, 2023). Under these conditions, real-time monitoring of stress-related physiological states may support human performance, safety, and adaptation. In ICE environments, human-agent teaming (HAT) offers a promising approach for real-time stress monitoring by integrating physiological data (e.g., HRV) with artificial intelligence (AI) and wearable devices (Giguère, Marois *et al.*, 2025). HAT systems hold the potential to be used to flag a high level of stress in an individual using physiological markers to detect stress and support adaptive interventions. However, it remains unclear whether it is feasible to infer perceived stress solely from physiological data, particularly in ICE environments.

The present study leverages two complementary experimental contexts: a controlled laboratory environment in which both physiological data and perceived stress ratings were collected (SPACE dataset; Giguère, Benesch *et al.*, 2025a), and an ICE mission in which only physiological measurements were available (DeepTime II experiment). We trained supervised models to associate HRV-derived features with perceived stress in the laboratory and then applied them to the ICE dataset. Rather than replacing subjective stress assessments, this approach examines whether models trained in the laboratory produce physiologically coherent patterns when used in an ICE environment without self-reports.

RESEARCH DESIGN AND METHODS

Participants

The data used in this study were collected from two independent datasets obtained under different experimental conditions. The SPACE dataset involves data collected from 120 healthy adults in Québec (Québec, Canada), who were free of any known conditions affecting autonomic response, in a controlled laboratory environment (Giguère, Benesch, *et al.*, 2025). The DeepTime II experiment involved 17 healthy adults who participated in a 15-day isolation study conducted in the Lombrives Cave (Occitanie, France). The DeepTime II participants were healthy adults aged between 18 and 50 years with no self-reported medical, neurological, psychiatric disorders, or addiction, not receiving ongoing medical treatment, and not enrolled in another biomedical research protocol. The two datasets were selected for their complementary characteristics: the laboratory dataset provided paired physiological and self-reported stress measures, whereas the ICE dataset offered a naturalistic environment where subjective assessments were unavailable. The protocol was designed to investigate human adaptation in an ICE environment. Informed consent was obtained from all participants in the two experiments, and the study was approved by the Comité de Protection des Personnes Est III 2025-A01059-40, France, for DeepTime II: No. 1000-497442-33, and Université Laval for SPACE dataset: No. 2024-335/11-09-2024.

Materials

Physiological data were collected using wearable sensors. In the SPACE dataset, HRV data was recorded using a BioHarness Zephyr device. Although acquisition devices differed, identical preprocessing pipelines were applied to mitigate variability. Perceived stress was assessed in the SPACE dataset using the VAS-S (Lesage *et al.*, 2012). The scale ranges from 0 to 10, where 0 represents no stress and 10 represents extreme stress. For DeepTime II, HRV data was recorded using a Polar H10 harness and no subjective stress assessment was available.

Experimental Protocol

The SPACE experiment involved participants completing multitasking 5-minute scenarios to induce stress, using OpenMATB (Cegarra *et al.*, 2020). Physiological data was continuously recorded throughout the tasks. Participants reported their perceived stress levels via the VAS-S. A manipulation check confirmed that perceived stress varied significantly throughout the experiment as a function of stress manipulations (Nicole *et al.*, 2026).

In the DeepTime II mission, participants took part in a 15-day period of isolation within the cave. Due to the absence of natural light and external time, participants experienced a complete loss of temporal references. The cave maintained stable conditions, with an ambient temperature of around 10°C and 99% humidity levels. Physiological data were collected at baseline (prior to full isolation), after six sleep–wake cycles, and at the experiment's conclusion during a 10-minute recording period, consisting of 5 minutes standing followed by 5 minutes in a seated position.

Preprocessing

Both datasets underwent identical preprocessing, segmentation and feature extraction, with subjective labelling applied only to the SPACE dataset. This ensured comparability and enabled model transfer between controlled and extreme environments (see Figure 1). Outlier detection was performed using a multi-step procedure inspired by Tiwari *et al.* (2020). Segmentation employed a 60-second sliding window with 75% overlap, in line with studies demonstrating the reliability of features over recordings of 60 seconds and the reliability of six ultra-short HRV features (Castaldo *et al.*, 2019).

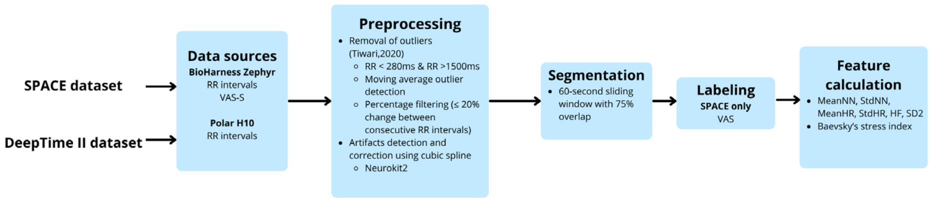

Figure 1: Overview of the data preparation pipeline applied to both datasets prior to model training.

BSI was not included as an input feature within the SPACE model. This decision was made to avoid redundancy with the HRV feature set and to preserve BSI as an independent physiological marker for comparative and contextual analyses between the two datasets. In DeepTime II, BSI was computed on each 60-second RR segment, consistent with the temporal resolution used for HRV feature extraction. All segments were retained for analysis without aggregation at the assessment time-point level.

Preliminary Analyses

Preliminary analyses were conducted at two levels. At the segment level, each 60-second time window was analysed separately for exploratory purposes. At the participant level, one observation per participant was used to ensure independence. For comparisons between datasets, segment-level BSI values were summarized using participant-specific medians. In the SPACE dataset, the association between perceived stress (VAS-S) and BSI was examined using participant-level mean aggregated values (mean VAS-S and mean BSI per participant) and Spearman correlation.

Modelling Approaches

Supervised learning, a technique that uses labeled data to train AI, was adopted to model the relationship between HRV-derived features and perceived stress, using the annotated SPACE dataset. Class imbalance was dealt with using Adaptive Synthetic (ADASYN) algorithm (He *et al.*, 2008), class-weighting as well as down-sampling the majority class in the three-class and two-class classifications. Both datasets were prepared using identical techniques. At each cross-validation iteration, models were trained on data from a subset of participants and validated on a disjoint subset of participants, ensuring that no individual appeared in both sets. Hyperparameters were optimized using randomized search with grouped cross-validation. The top-performing models based on balanced accuracy were applied to the DeepTime II dataset without retraining to generate physiology-based stress proxies. Analyses on DeepTime II focused on descriptive and association-based analyses, as no self-reported stress labels were available.

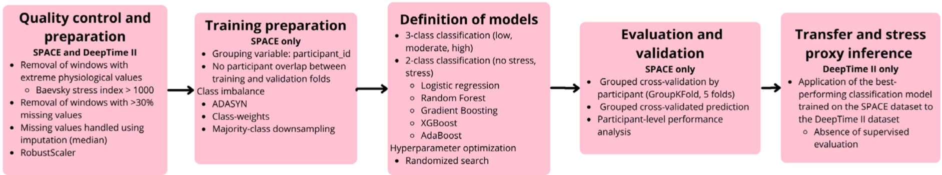

Figure 2: Overview of the training, validation and transfer pipeline to infer perceived stress from HRV.

RESULTS

The results are divided into three sections: initial analyses of label distribution and physiological differences, evaluation of model performance on the SPACE dataset for classification, and examination of model transfer to the DeepTime II dataset through distributional and association-based analyses.

Preliminary Analyses

As shown in Figure 3, VAS-S labels were unevenly distributed in the SPACE dataset, with a higher proportion of low stress ratings and progressively fewer observations at higher stress levels.

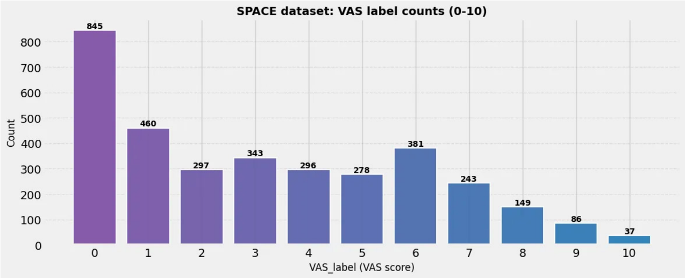

Figure 3: Distribution of VAS-S Scores across the SPACE dataset.

Mean VAS-S scores in the SPACE dataset differed across scenario types, $F(3.34, 407.93) = 100.41$, $p < 0.001$, Greenhouse-Geisser corrected, with baseline scenarios showing the lowest stress levels and stress-inducing scenarios the highest (*p*s < .001; see also Figure 4). These differences support the construct validity of the VAS-S as a label for further classification using supervised learning.

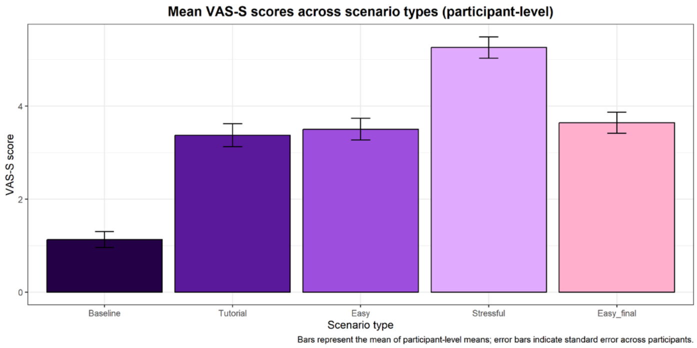

Figure 4: Participant-level mean VAS-S scores across scenario types in SPACE.

BSI values were higher in the DeepTime II dataset than in the controlled laboratory dataset. At the participant level, median BSI values were significantly higher in DeepTime II compared to SPACE ($U = 638$, $p = .02$, $r_r = .20$), indicating increased autonomic strain in the DeepTime II. In the SPACE dataset, no significant association was observed between VAS-S and BSI at the participant level. For each participant, VAS-S and BSI were averaged within each participant. Spearman's correlation across participants revealed no significant relation ($r_S = -.07$, $p = .47$).

Classification Results on SPACE

Data were segmented into 60-second windows with 75% overlap for analysis. The SPACE dataset was labeled for multi-class classification into three stress categories based on VAS-S scores: low stress (0-2, $n = 1{,}565$), moderate stress (3-6, $n = 1{,}275$), and high stress (7-10, $n = 503$; see Ritvanen *et al.*, 2006; Rodrigues *et al.*, 2018). For binary classification, scores from 0 to 2 were labeled as no stress ($n = 1{,}565$), while scores from 3 to 10 were labeled as stress (n = 1,778). Supervised learning models were trained to evaluate how well cardiac autonomic markers could reflect perceived stress.

Multi-class classification. For the three-class classification task, the best-performing model was a XGBoost classifier combined with ADASYN. Using grouped cross-validation by participant on the SPACE dataset, this model achieved a balanced accuracy of 37.7% and an overall accuracy of 44.7%, with a macro-averaged F1-score of 37.8% (chance classification = 33.3%). Specifically, the low stress class showed the highest performance, with an F1-score of 52.6%, a precision of 50.4%, and a recall of 54.9%. For the moderate stress class, the F1-score was 43.1%, with a precision of 42.1% and a recall of 44.2%. Performance for the high stress class was lower, with an F1-score of 17.7%, a precision of 23.7%, and a recall of 14.1%.

Comparable performance was observed across models using different class imbalance handling strategies. The balanced accuracy values for the down-sampling approach, ADASYN, class-weighted, and non-rebalanced models ranged from approximately 33.7% to 37.7%. At the participant level, multi-class classification performance showed substantial inter-individual variability. Balanced accuracy values ranged from near 0% to approximately 76.7%, while accuracy varied from near 0% to approximately 75.8% across participants. A large proportion of participants did not exhibit all three stress classes, resulting in undefined multi-class metrics for these individuals.

Binary classification. For the binary classification task, the model that performed best was logistic regression with no class imbalance handling. The model achieved a balanced accuracy of 55.5% and an overall accuracy of 56.4% (chance classification = 50%). The F1-score for the stress class was 63.0%, with a precision of 57.4% and a recall of 69.7%. For the no-stress class, F1, precision and recall were 47.0%, 54.5%, and 41.3%, respectively. Comparable performance was achieved across models using different class imbalance handling strategies, including class weighting, majority-class down-sampling, ADASYN and no rebalancing technique. Balanced accuracy values consistently ranged from 53.6% to 56.4%. No single rebalancing approach was found to be superior. Classification performance among participants exhibited significant variability, with balanced accuracy ranging from around 35% to a high of 94.0%. Individual accuracy varied between 30.6% and 90.0%. For those with only one stress class, participant-level metrics could not be calculated.

Transfer to DeepTime II

The best-performing classification models trained on the SPACE dataset were applied to physiological data from all three DeepTime II assessment points (prior to isolation, mid-isolation, and end of confinement) to generate physiology-based stress proxies in the absence of self-reported stress measures. The absence of VAS-S labels meant that no supervised performance evaluation could be conducted for DeepTime II.

When applied to the DeepTime II dataset, predicted stress classes (VAS proxy) were examined in relation to BSI. At the segment level, median BSI values differed descriptively across predicted classes (low: *Med* = 190.0, IQR = [122.4–313.4]; moderate: *Med* = 116.8, IQR = [57.9–215.7]; high: *Med* = 233.5, IQR = [89.4–420.9]). This pattern persisted at the participant level, where median BSI values were higher for participants classified as high stress compared to moderate and low stress (low: *Med* = 197.8, IQR = [147.8–231.2]; moderate: *Med* = 137.7, IQR = [73.5–188.5]; high: *Med* = 284.8, IQR = [109.3–379.8]). This pattern suggests internal physiological coherence between predicted classes and autonomic activation in the absence of self-reported measures. However, because both the classification model and BSI are derived from HRV features, this association should not be interpreted as an independent validation of perceived stress inference. Substantial overlap in BSI values was observed across predicted VAS proxy categories, highlighting the continuous nature of autonomic stress responses and the limitations of discrete classification boundaries in extreme environments.

DISCUSSION

This study examined whether perceived stress can be inferred from autonomic markers and whether AI models trained in a controlled laboratory context can be applied to an ICE environment. Specifically, HRV-derived features were used in a laboratory to train supervised classification models in a labeled lab dataset and subsequently applied to physiological data collected during the DeepTime II experiment.

The preliminary analyses revealed that, although perceived stress varied reliably across experimental conditions in the SPACE dataset, its association with autonomic strain, as indexed by the BSI, was weak at both the segment and participant levels. This weak association suggests that perceived stress and autonomic activation are not strongly aligned, even in a controlled laboratory setting. Although perceived stress is generally expected to elicit physiological responses, contemporary conceptual frameworks emphasize that subjective appraisal and physiological activation represent partially distinct dimensions of the stress response (Epel *et al.*, 2018). This dissociation may partly explain the modest performance of the HRV-based models, which reflects the multidimensional nature and subjectivity of stress.

In the SPACE dataset, HRV-based models trained on laboratory data were only modestly above chance level in both binary (50%) and three-class (33.3%) classification tasks. The choice of rebalancing technique did not appear to affect the performance of the models. The modest performance and high inter-individual variability in SPACE likely reflect intrinsic limitations of generalized HRV-based stress models, suggesting that relative within-individual autonomic changes may be more informative than absolute stress predictions in extreme environments. Self-perceived stress is a broad and multidimensional concept that does not always translate directly onto biological stress markers, as stressed within recent stress frameworks (Epel *et al.*, 2018). Furthermore, relative autonomic alterations within an individual may provide more information than absolute stress predictions across people. Strong inter-individual variability was observed, suggesting the potential relevance of individualized modelling approaches. However, such approaches could not be explored in the present study due to insufficient data per participant. Also, stress inference based on autonomic markers (e.g., BSI) may be better suited for tracking relative changes within individuals over time rather than predicting absolute stress levels across individuals.

When applied to the DeepTime II dataset, the best-performing models produced stress proxies that showed variation in BSI across predicted classes, despite the absence of self-reported stress labels. These findings suggest that cardiac autonomic markers provide limited proxies of perceived stress when subjective reports are unavailable, a common constraint in extreme and isolated environments. In this dataset, the relationship between self-perceived stress (VAS-S) and HRV features appeared weak and highly variable across individuals. However, the moderate performance and inter-individual variability observed in the controlled setting highlight the limitations of generalized models and the need for more individualized or context-aware approaches.

Previous work by Castaldo *et al.* informed the selection of HRV features extracted from the SPACE dataset by demonstrating that several cardiac autonomic markers remain reliable when computed over 60-second windows, as opposed to the conventional 5-minute segments (Castaldo *et al.*, 2019). In their study, an accuracy of 88% was reported when discriminating between rest and experimentally induced stress under controlled conditions (Castaldo *et al.*, 2019). In contrast, the present study adopted participant-wise validation to specifically assess inter-individual generalization rather than optimize laboratory performance. This more restrictive evaluation framework highlights the challenges faced by generalized HRV-based stress models when applied to novel individuals and real-world deployment scenarios. This evaluation strategy was necessary to ensure compatibility with the DeepTime II setting, where subject-dependent calibration and self-reported stress measures were unavailable.

In HAT settings, stress inference systems may be more effective if, rather than estimating absolute stress levels, they monitor deviations from an individual's baseline. Such a system could facilitate the early detection of stress-related changes and provide adaptive support in ICE environments, where subjective reports are often unavailable. This study reinforces that HRV features alone are insufficient for reliable stress detection (Martinez *et al.*, 2022), and that integrating subjective stress measures is considered the best practice (Vaessen *et al.*, 2021).

The study evaluates the robustness of generalized HRV-based stress models across different contexts, emphasizing the limitations of using autonomic markers for stress inference without retraining or subjective labels. It suggests that such HRV-based models may have inherent constraints when predicting subjective stress across individuals and environments.

Limitations

This study focused solely on RR-derived HRV features and did not incorporate contextual, behavioral, or multimodal physiological data, which may have limited the ability of the models to capture perceived stress. While some participants in the SPACE dataset had sufficient labelled segments for individualized model exploration, the significantly smaller number of segments per participant in the DeepTime II dataset, coupled with the absence of subjective stress labels, hindered model development across environments. Furthermore, even with harmonized preprocessing, variations in acquisition devices and environmental circumstances among datasets would have produced residual variability, especially in terms of data quality.

CONCLUSION

This study evaluated the robustness of generalized HRV-based stress models across laboratory and extreme ICE contexts. Performance remained modest and highly variable across individuals, and stress proxies derived in the ICE setting showed physiological coherence but substantial overlap across categories. These findings highlight the practical limits of HRV-only stress modelling and support the need for individualized and multimodal approaches in extreme environments.

ACKNOWLEDGMENT

The authors would like to acknowledge Olivier Morin, Laurie Dionne, Jeanne Nicole, Laurent Berthod, Danielle Benesch, Sébastien Tremblay, Denis Ouellet and Tanya S. Paul for their participation in the data collection for the SPACE dataset. The authors would also like to acknowledge the Human Adaptation Institute for providing the DeepTime II dataset, as well as Lubin Le Mélinaire for his role in this project. Raphaëlle Giguère holds a research grant for her doctoral studies from the VITAM Research Center and Canadian Space Agency.

REFERENCES

Arbeille, P., Zuj, K.A., Besnard, S., Mauvieux, B., Hingrand, C., Delaunay, P.-L., Hughson, R.L. and Clot, C. (2023) "Ultrasound assessments of organs and blood vessels before and after 40 days isolation in a cavern (Deep Time experiment 2021)", *Frontiers in Physiology*, 14. Available at: https://doi.org/10.3389/fphys.2023.1174565

Baevsky, R. and Chernikova, A. (2017) "Heart rate variability analysis: physiological foundations and main methods", *Cardiometry*, 10, pp. 66–76. Available at: https://doi.org/10.12710/cardiometry.2017.10.6676

Bernard, C. (1957) *An introduction to the study of experimental medicine*. New York: Dover Publications.

Castaldo, R., Montesinos, L., Melillo, P., James, C. and Pecchia, L. (2019) "Ultra-short term HRV features as surrogates of short term HRV: a case study on mental stress detection in real life", *BMC Medical Informatics and Decision Making*, 19(1), p. 12. Available at: https://doi.org/10.1186/s12911-019-0742-y

Cegarra, J., Valéry, B., Avril, E., Calmettes, C. and Navarro, J. (2020) "OpenMATB: a multi-attribute task battery promoting task customization, software extensibility and experiment replicability", *Behavior Research Methods*, 52(5), pp. 1980–1990. Available at: https://doi.org/10.3758/s13428-020-01364-w

Curtis, A.L., Lechner, S.M., Pavcovich, L.A. and Valentino, R.J. (1997) "Activation of the locus coeruleus noradrenergic system by intracoerulear microinfusion of corticotropin-releasing factor: effects on discharge rate, cortical norepinephrine levels and cortical electroencephalographic activity", *The Journal of Pharmacology and Experimental Therapeutics*, 281(1), pp. 163–172.

Epel, E.S., Crosswell, A.D., Mayer, S.E., Prather, A.A., Slavich, G.M., Puterman, E. and Mendes, W.B. (2018) "More than a feeling: a unified view of stress measurement for population science", *Frontiers in Neuroendocrinology*, 49, pp. 146–169. Available at: https://doi.org/10.1016/j.yfrne.2018.03.001

Giguère, R., Benesch, D., Paul, T.S., Ouellet, D., Chamberland, C., Marois, A. and Tremblay, S. (2025) "*SPACE – Stress and Physiological Assessment in a Controlled Environment*". [dataset]. Zenodo. Available at: https://doi.org/10.5281/zenodo.15330494 (Accessed January 10, 2026).

Giguère, R., Marois, A., Fortin-Guichard, D., Charest, J., Rocheleau, J., Thompson, K., Stolberg, M., St-Martin, P., Bergouignan, A., Niaussat, V., Paquette, J.-S., Gagnon, M.-P., Sasseville, M. and Rhéaume, C. (2025) "Space exploration and lifestyle medicine: a narrative review of its implications for astronaut health and remote Earth-based environments", *Frontiers in Physiology*, 16. Available at: https://doi.org/10.3389/fphys.2025.1729596

He, H., Bai, Y., Garcia, E.A. and Li, S. (2008) "ADASYN: adaptive synthetic sampling approach for imbalanced learning", in *2008 IEEE International Joint Conference on Neural Networks*. Piscataway, NJ: IEEE, pp. 1322–1328.

Kim, H.-G., Cheon, E.-J., Bai, D.-S., Lee, Y.H. and Koo, B.-H. (2018) "Stress and heart rate variability: a meta-analysis and review of the literature", *Psychiatry Investigation*, 15(3), pp. 235–245. Available at: https://doi.org/10.30773/pi.2017.08.17

Le Roy, B., Martin-Krumm, C., Pinol, N., Dutheil, F. and Trousselard, M. (2023) "Human challenges to adaptation to extreme professional environments: a systematic review", *Neuroscience & Biobehavioral Reviews*, 146, p. 105054. Available at: https://doi.org/10.1016/j.neubiorev.2023.105054

Lesage, F.-X., Berjot, S. and Deschamps, F. (2012) "Clinical stress assessment using a visual analogue scale", *Occupational Medicine*, 62(8), pp. 600–605. Available at: https://doi.org/10.1093/occmed/kqs140

Manzey, D., Lorenz, B. and Poljakov, V. (1998) "Mental performance in extreme environments: results from a performance monitoring study during a 438-day spaceflight", *Ergonomics*, 41(4), pp. 537–559. Available at: https://doi.org/10.1080/001401398186991

Marques, A.H., Silverman, M.N. and Sternberg, E.M. (2010) "Evaluation of stress systems by applying noninvasive methodologies: measurements of neuroimmune biomarkers in the sweat, heart rate variability and salivary cortisol", *Neuroimmunomodulation*, 17(3), pp. 205–208. Available at: https://doi.org/10.1159/000258725

Martinez, G.J., Grover, T., Mattingly, S.M., Mark, G., D'Mello, S., Aledavood, T., Akbar, F. and Striegel, A. (2022) "Alignment between heart rate variability from fitness trackers and perceived stress: perspectives from a large-scale in situ longitudinal study of information workers", *JMIR Human Factors*, 9(3), p. e33754. Available at: https://doi.org/10.2196/33754

Nicole, J., Giguère, R., Ouellet, D., Benesch, D., Tremblay, S. and Marois, A. (2026) "Validating stress induction for spaceflight scenarios: a manipulation check across acute stressor types in simulated task environments", in *Proceedings of the 6th International Conference on Human Interaction & Emerging Technologies: Artificial Intelligence & Future Applications (IHIET-AI 2026)*. València: Universitat Politècnica de València. DOI : https://doi.org/10.54941/AHFE-Paper-0123

Rajendra Acharya, U., Paul Joseph, K., Kannathal, N., Lim, C.M. and Suri, J.S. (2006) "Heart rate variability: a review", *Medical & Biological Engineering & Computing*, 44(12), pp. 1031–1051. Available at: https://doi.org/10.1007/s11517-006-0119-0

Ritvanen, T., Louhevaara, V., Helin, P., Väisänen, S. and Hänninen, O. (2006) "Responses of the autonomic nervous system during periods of perceived high and low work stress in younger and older female teachers", *Applied Ergonomics*, 37(3), pp. 311–318. Available at: https://doi.org/10.1016/j.apergo.2005.06.013

Rodrigues, S., Paiva, J.S., Dias, D. and Cunha, J.P.S. (2018) "Stress among on-duty firefighters: an ambulatory assessment study", *PeerJ*, 6, p. e5967. Available at: https://doi.org/10.7717/peerj.5967

Rotenberg, S. and McGrath, J.J. (2016) "Inter-relation between autonomic and HPA axis activity in children and adolescents", *Biological Psychology*, 117, pp. 16–25. Available at: https://doi.org/10.1016/j.biopsycho.2016.01.015

Selye, H. (1956) "Endocrine reactions during stress", *Current Researches in Anesthesia & Analgesia*, 35(3), pp. 182–193.

Selye, H. (1978) *The stress of life*. Rev. edn. New York: McGraw-Hill.