DOI: 10.4274/anatoljmed.2026.09226
Anatol J Gen Med Res 2026;36(2):132-41

CLINICAL RESEARCH / KLİNİK ARAŞTIRMA

ANATOLIAN JOURNAL of GENERAL MEDICAL RESEARCH

Integrative Analysis of Stress During Academic Exams Using Wearable Sensor Data

Akademik Sınavlarda Stres Tepkilerinin Giyilebilir Sensör Verileriyle Bütüncül Analizi

Hasan Kazdağlı[1], Hasan Fehmi Özel[2]

[1]İzmir University of Economics Vocational School of Health Services, Department of Medical Services and Techniques, İzmir, Türkiye
[2]Manisa Celal Bayar University Vocational School of Health Services, Department of Medical Services and Techniques, Manisa, Türkiye

Cite as: Kazdağlı H, Özel HF. Integrative analysis of stress during academic exams using wearable sensor data. Anatol J Gen Med Res. 2026;36(2):132-41

Abstract

Objective: To comprehensively analyze students' physiological stress responses and their relationship with academic performance by using heart rate variability (HRV) and electrodermal activity (EDA) data collected via wearable sensors during academic examinations.

Methods: This study utilized the publicly available wearable exam stress dataset from PhysioNet. HRV and EDA data were recorded using Empatica E4 wristbands worn by ten university students during three exam sessions (midterm 1, midterm 2, and final). HRV parameters [low frequency (LF), high frequency (HF), LF/HF] and the tonic and phasic components of EDA were extracted. Differences between sessions were analyzed using the Mann-Whitney U test, while correlations with exam grades were evaluated using Spearman's rank correlation coefficient. In addition, individual variability and the relationship between HRV and EDA were examined on a session-by-session basis.

Results: Among HRV parameters, LF power ($p=-0.440$, $p=0.015$) and LF/HF ratio ($p=-0.363$, $p=0.048$) showed significant negative correlations with academic performance. Mean EDA values did not significantly differ across sessions ($p>0.7$); however, phasic components indicated increased arousal during the midterm 2 session. Correlations between EDA and grades were negative in midterm 1 ($p=-0.552$) and moderately positive in the final ($p=0.406$). Combined analysis of HRV and EDA data revealed concurrent autonomic dysregulation and acute stress responses, particularly during the midterm 2 session.

Conclusion: These findings indicate that academic examinations are associated with heterogeneous autonomic nervous system response patterns among students. The integrated evaluation of HRV and EDA contributes to a more comprehensive characterization of physiological stress responses at both the group and individual levels. Wearable sensor-derived physiological data may offer a useful framework for informing future research on stress monitoring and the development of individualized support approaches in academic settings.

Keywords: Heart rate variability, electrodermal activity, exam stress, integrative analysis

Öz

Amaç: Bu çalışmanın amacı, akademik sınavlar sırasında giyilebilir sensörlerle elde edilen kalp hızı değişkenliği (HRV) ve elektrodermal aktivite (EDA) verilerini kullanarak öğrencilerin fizyolojik stres tepkilerini ve bu tepkilerin akademik başarıyla ilişkisini bütüncül olarak analiz etmektir.

Yöntem: Bu çalışmada, PhysioNet'te açık erişimli olarak sunulan *wearable exam stress dataset* kullanıldı. On üniversite öğrencisinin üç sınav oturumu (ara sınav[1], ara sınav[2] ve final) sırasında giyilebilir Empatica E4 bilekliği ile kaydedilen HRV ve EDA verileri analiz edildi. HRV parametreleri [düşük frekans (LF),

Address for Correspondence/Yazışma Adresi: Asst. Prof., Hasan Kazdağlı, İzmir University of Economics Vocational School of Health Services, Department of Medical Services and Techniques, İzmir, Türkiye
E-mail: kazdaglihasan@gmail.com
ORCID ID: orcid.org/0000-0001-6617-604X

Received/Geliş tarihi: 27.11.2025
Accepted/Kabul tarihi: 29.01.2026
Published date/Yayınlanma tarihi: 31.08.2026

Öz

yüksek frekans (HF), LF/HF] ve EDA'nın tonik-fazik bileşenleri çıkarıldı; oturumlar arası farklar Mann-Whitney U testi ile, notlarla ilişkiler ise Spearman korelasyonu ile değerlendirildi. Ayrıca bireysel değişkenlik ve HRV-EDA ilişkileri de oturum bazında incelendi.

Bulgular: HRV parametrelerinden LF gücü (p=-0,440, p=0,015) ve LF/HF oranı (p=-0,363, p=0,048) akademik başarıyla anlamlı düzeyde negatif korelasyon gösterdi. EDA ortalama değerleri oturumlar arasında anlamlı farklılık göstermedi (p>0,7); ancak fazik bileşenler ara sınav 2 oturumunda artmış uyarılma düzeylerini yansıttı. EDA ile notlar arasında ara sınav 1'de negatif (p=-0,552) ve final oturumunda pozitif yönde orta düzeyde (p=0,406) korelasyon gözlendi. HRV ve EDA verilerinin birlikte incelenmesi, özellikle ara sınav 2'de otonom düzensizlik ile ani stres tepkilerinin birlikte seyrettiğini gösterdi.

Sonuç: Bu bulgular, akademik sınavların öğrencilerde heterojen otonom sinir sistemi yanıtlarıyla ilişkili olduğunu göstermektedir. HRV ve EDA'nın bütüncül olarak değerlendirilmesi, stres tepkilerinin hem grup hem de bireysel düzeyde daha kapsamlı biçimde tanımlanmasına katkı sağlamaktadır. Giyilebilir sensörlerden elde edilen fizyolojik veriler, akademik ortamlarda stres izlenmesine yönelik gelecekteki araştırmalar ve bireyselleştirilmiş destek yaklaşımlarının geliştirilmesi için yararlı bir çerçeve sunabilir.

Anahtar Kelimeler: Kalp atım hızı değişkenliği, elektrodermal aktivite, sınav stresi, bütüncül analiz

Introduction

Exam stress is common among students at various academic levels and typically results from a complex interaction of psychological, physiological, and behavioral factors[1]. Understanding the impact of exam stress on students requires a multifaceted approach, as it significantly influences both their mental well-being and academic performance[2]. Although exam stress is considered a natural part of academic life, it is closely linked to students' stress perceptions, coping strategies, and ultimately their academic achievement[3].

The impact of stress on academic achievement is a topic that has been examined by numerous researchers. The stress experienced by students leads to negative outcomes such as decreased exam performance and reduced course attendance[4]. Similarly, students who experience high levels of stress often suffer from fatigue, which creates a cycle between declining performance and increasing stress[5]. The widespread effects of stress are not limited to academic outcomes; by influencing behavior, it can lead to maladaptive coping mechanisms such as substance use and social withdrawal, which further exacerbate academic difficulties[6,7].

Students who successfully develop effective ways of coping with stress tend to show improved academic performance, along with reduced habits such as substance and tobacco use[8]. Therefore, early identification of students who are prone to stress and providing them with effective stress-management skills are of great importance[9]. Today, the most commonly used method for assessing exam stress consists of questionnaire-based studies in which students evaluate themselves[10]. Many studies have shown that these scales may present limitations in terms of reliability and validity, which raises questions about their usefulness in accurately measuring exam-related stress[11].

First, the development of stress measurement scales often involves multiple psychological constructs, which can lead to inconsistencies in practice. For example, although the beliefs about stress scale has demonstrated adequate validity when administered before and after major exams, the reliability of such scales may vary depending on contextual factors such as the demographic characteristics of the student sample[12]. This indicates that although some scales may provide reliable results under specific conditions, they are not universally applicable across different academic settings or types of stress, which weakens overall assessment reliability.

Moreover, Zunhammer et al.[13] have shown that certain psychological traits can significantly distort the measurements obtained from stress scales administered during examinations. This study demonstrated that some personality characteristics are not sufficiently accounted for in traditional stress assessment models, which may negatively affect the reliability and interpretability of the results. The presence of heightened somatic symptoms associated with exam stress, alongside psychological factors, makes the direct interpretation of scores obtained from stress measurement tools more complex and raises concerns about their reliability across different student groups[14].

Research shows that academic stress significantly affects various physiological markers, including cortisol levels, heart rate variability (HRV), and electrodermal activity (EDA)[3,15].

HRV is widely used as a valuable biomarker for evaluating stress responses, particularly in high-pressure environments such as exams[16]. The relationship between HRV and academic stress has been examined in various studies. For example, Obi et al.[17] reported that students exhibited lower HRV values before and during exams, which was interpreted as an indicator of heightened mental stress. This reduced HRV observed during exam periods has been associated with increased sympathetic dominance and decreased parasympathetic activity. Similarly, Hammoud et al.[18] reported a decrease in students' HRV levels during exams, with a significant increase observed after exams. This finding indicates a reduction in stress following the exam period.

EDA, also commonly known as galvanic skin response, is an important physiological measure that reflects sympathetic nervous system activity[19]. This measurement has various applications in evaluating stress, particularly in educational or competitive environments such as exams[19]. EDA stands out as an effective and non-invasive indicator of stress because it correlates with sympathetic nervous system arousal, which varies with stress levels[20]. EDA measurements can evaluate emotional and stress-related states without requiring complex equipment, and this method is considered ideal for educational settings where exam stress is common[21]. Furthermore, studies have shown that stress can lead to a decline in cognitive performance and an increase in anxiety levels. This is particularly evident in high-pressure situations such as exams, and physiological responses like EDA can provide real-time feedback on students' stress levels[22,23].

The study conducted by Rahma et al.[24] demonstrated that both the tonic (baseline) and phasic (stimulus-responsive) components of EDA are effective in distinguishing different levels of cognitive and emotional stress. This finding indicates that EDA enables the monitoring of stress responses during critical moments such as exams[24].

However, studies that analyze EDA and HRV data together and relate them to students' academic performance remain limited. Therefore, in our study we aimed to examine the relationship between physiological stress responses and academic performance by comprehensively evaluating EDA and HRV data collected with wearable sensors during academic exams.

Accordingly, the present study adopts an exploratory and hypothesis-generating approach, aiming to characterize integrated autonomic stress responses during real-world academic examinations rather than to establish definitive causal relationships. By leveraging an ecologically valid wearable dataset, this work seeks to identify physiological patterns that may inform future large-scale and controlled investigations in educational psychophysiology.

Materials and Methods

Study Design and Dataset

This study was conducted using the "wearable exam stress dataset", an open-access data set[25]. The dataset contains electrophysiological signals collected from ten university students during three different academic exam sessions (midterm 1, midterm 2, and final), using the Empatica E4 wearable wrist device[26]. Heartbeat intervals [inter-beat interval (IBI)], EDA, and body temperature data were continuously recorded from participants using the devices.

The dataset does not include resting or non-exam baseline physiological recordings. Therefore, all analyses were conducted using within-session and between-session comparisons across examination periods, rather than deviations from individual baseline autonomic states.

Ethical Aspect of the Research

Our study was conducted with the approval of the İzmir University of Economics Health Sciences Research Ethics Committee (approval no: E-97429853-050.04-111205, date: 05.11.2025). The dataset is available to researchers via the PhysioNet platform, and because participant information is anonymized, an informed consent form is not required. In our study, no direct intervention or interaction was carried out, as the dataset was used solely for analysis. Therefore, the research qualifies as a secondary data analysis.

Participants

Data from 10 university students in the wearable exam stress dataset were analyzed. The participants were healthy individuals with no psychiatric or neurological diagnoses and were included in the study on a voluntary basis[25]. In the dataset, participants' identities were kept confidential, and only limited demographic information, such as age, sex, and academic performance, was provided. Participants included male and female students, and the reported age range was 20-25 years[25]. For each participant, electrophysiological data from three exam sessions (midterm 1, midterm 2, and final) and the students' exam scores were obtained.

Data Collection

Data were collected using the Empatica E4 wrist-worn device during participants' exam sessions. This wearable sensor can simultaneously record various physiological parameters. In addition to the two primary electrophysiological signals, IBI and EDA, body temperature was also recorded within the scope of the study[24]. Figure 1 illustrates a representative IBI tachogram derived from wearable photoplethysmography (PPG) recordings.

IBI data obtained using an optical PPG sensor were used for HRV calculations. EDA data were continuously recorded using two electrodes placed on the inner surface of the wrist and evaluated as indicators of sympathetic nervous system activity. Body temperature measurements were continuously monitored via the temperature sensor located on the device's body and were analyzed for the time periods corresponding to the exam sessions[24]. Accordingly, skin temperature should be viewed as a complementary, context-dependent marker that gains interpretive value primarily when integrated with HRV and EDA indices.

For each participant, nearly 24 hours of data were recorded throughout the exam day[24]. Because of the time-stamped nature of the data, segments corresponding to the periods immediately before and during the exam were extracted and preprocessed for analysis. Noise, motion artifacts, and missing signals were removed using appropriate filtering and cleaning methods prior to data analysis.

Although the Empatica E4 device provides validated PPG-derived IBI measurements, it is acknowledged that wrist-based PPG signals are more susceptible to motion artifacts and may exhibit reduced precision compared to electrocardiography-based recordings. Consequently, HRV results should be interpreted with appropriate caution, particularly in high-movement contexts.

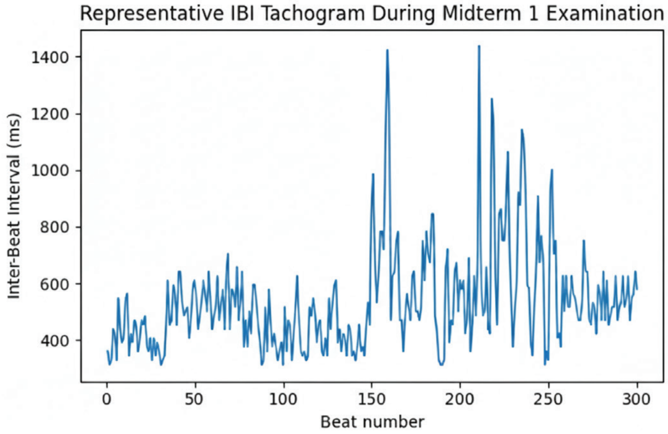

Figure 1. Representative inter-beat interval (IBI) tachogram derived from wearable photoplethysmography recordings during the midterm 1 examination. The tachogram, after artifact cleaning and preprocessing, illustrates beat-to-beat variability and demonstrates the signal quality used for subsequent HRV and non-linear analyses

HRV: Heart rate variability

HRV and EDA Analyses

IBI data obtained from the PPG sensor of the Empatica E4 device were used for HRV analysis. Raw IBI series were initially subjected to automated artifact detection and cleaning procedures to identify ectopic beats, motion-related disturbances, and missing values. Artifact-contaminated intervals were corrected using interpolation following sequence verification, and only segments with adequate signal quality and a minimum duration of 5 minutes were retained for subsequent analyses. Because the dataset did not include resting or non-exam baseline recordings, HRV metrics were interpreted using within-session and between-session comparisons across examination periods, rather than as deviations from individual normative autonomic values.

HRV analysis included time-domain, frequency-domain, and non-linear parameters. For frequency-domain analyses, low frequency [(LF): 0.04-0.15 Hz] and high frequency [(HF): 0.15-0.4 Hz] components, as well as the LF/HF ratio, were calculated[3]. These indices reflect sympathetic modulation, parasympathetic activity, and overall autonomic balance, respectively[27]. For frequency analysis, power spectral density was estimated using the Welch method[28]. HRV analysis was performed using the open-source Python libraries NeuroKit2 and HRV-analysis. HRV metrics were computed using open-source Python libraries NeuroKit2 and HRV-analysis, employing consistent preprocessing and parameter settings across all exam sessions to ensure methodological standardization. HRV parameters were calculated separately for each examination session and subsequently evaluated both across individuals and between exam types. In addition to conventional HRV indices, non-linear measures derived from detrended fluctuation analysis (DFA) were included. Specifically, the dS1, dS2, and TDs parameters, proposed by Ozel and Kazdagli[29] as quantitative indices of short-term, long-term, and total physiological complexity loss in heart rate time series, were computed to characterize stress-related alterations in autonomic complexity. Given the known limitations of wrist-based PPG-derived HRV measurements, particularly their susceptibility to motion artifacts, all results were interpreted with appropriate caution.

EDA signals were analyzed for the time intervals corresponding to each examination session. Raw EDA signals, sampled at 4 Hz, were initially inspected for signal dropouts and motion-related artifacts. Noise reduction was performed using low-pass filtering, and segments with excessive artifacts were excluded to ensure signal reliability. Because environmental and individual factors influencing EDA could not be experimentally controlled, all analyses were conducted using a consistent preprocessing pipeline across sessions to maintain methodological standardization.

Following preprocessing, EDA signals were decomposed into tonic (skin conductance level) and phasic [skin conductance response, (SCR)] components. The tonic component reflects the overall level of sympathetic arousal, whereas the phasic component captures rapid, stimulus-related sympathetic activations[24]. Decomposition was performed using the convex-optimization-based cvxEDA algorithm with default regularization parameters as implemented in NeuroKit2.

For each participant and examination session, summary metrics including mean EDA level, number of SCRs, and maximum SCR amplitude were calculated. All EDA analyses were conducted in Python using the NeuroKit2 and BioSPPy libraries, with identical parameter settings applied across sessions to ensure analytical consistency.

Academic Performance Data

To evaluate the participants' academic performance, individual exam scores from each session (midterm 1, midterm 2, and final) were used. These scores were provided to researchers as part of the "wearable exam stress dataset", with midterm 1 and midterm 2 graded on a 100-point scale. Final exam scores were normalized to a 100-point scale before being used in the analyses.

Statistical Analysis

Statistical analysis of the data was performed using GraphPad Prism (version 10.4.0). Data were expressed as medians. A significance level of $p<0.05$ was accepted for all analyses. To evaluate differences in HRV and EDA parameters across exam sessions (midterm 1, midterm 2, and final), the non-parametric Mann-Whitney U test was applied.

Spearman's rank correlation coefficient (ρ) was calculated to determine the relationships between HRV and EDA parameters and exam scores. In addition, individual physiological response profiles were examined, and the relationships between HRV and EDA components were evaluated on a session-by-session basis.

Correlation plots present the correlation coefficients and their significance levels.

Exam score data were treated as continuous variables and analyzed at both the individual and session levels. Because the score distributions were not normal, non-parametric correlation analyses were preferred. The relationships between physiological data and academic performance were examined using Spearman's rank correlation. Moreover, to assess the impact of differing stress levels on performance, participants were grouped according to HRV and EDA parameters.

Differences in EDA and HRV parameters across the different exams were assessed using the Friedman test. For parameters that reached statistical significance ($p<0.05$), post-hoc analyses (Wilcoxon matched-pairs test with Bonferroni correction) were performed.

Results

Comparison of HRV, EDA, and Body Temperature Parameters Across Exams

The HRV and EDA metrics, together with the corresponding examination scores, are summarized in Table 1 and presented as median ± interquartile range (IQR). Both DFAα_1 and the stress index exhibited pronounced elevations during midterm 1 ($p=0.000265$ and $p=0.00145$, respectively), indicating an early-session perturbation in autonomic regulatory dynamics. Conversely, the tonic component of the EDA signal was significantly attenuated during the final examination ($p=0.00048$), suggesting diminished sympathetic arousal in later stages of the assessment sequence. No additional physiological parameters demonstrated session-dependent divergence ($p>0.05$), indicating relative stability in other autonomic indices across the exam period.

Within-participant analyses provided further evidence of session-specific modulation of autonomic function. Mean heart rate (Figure 2) reached its maximum during midterm 1 and subsequently declined, while DFAα_1 (Figure 3), a well-established marker of short-term fractal cardiac dynamics, demonstrated an analogous pattern of early-session elevation followed by progressive reduction. Friedman test results corroborated these effects (mean hazard ratio: $p=0.0001$; DFAα_1: $p=0.0498$). Post-hoc contrasts revealed significant decreases in mean heart rate between midterm 1 and midterm 2, and between midterm 1 and the final examination ($p=0.0117$ for both). A significant decline in DFAα_1 emerged exclusively between midterm 1 and the final

Table 1. HRV parameters, EDA values, and exam scores of students across the exams

Parameters	Midterm 1	Midterm 2	Final
Exam score (%)	77.50±14.75	77.00±22.50	76.50±19.62
Mean HR (bpm)	121.65±11.00	95.83±11.12	91.41±15.45
SDNN (ms)	30.35±25.81	31.88±13.06	27.38±7.55
RMSSD (ms)	16.65±10.77	22.71±7.78	21.86±1.83
Stress index	18.48±8.89	10.90±4.09	13.35±3.06
VLF (%)	74.14±27.94	69.13±15.06	71.85±10.16
LF (%)	45.41±21.66	46.28±12.34	50.95±10.18
HF (%)	36.51±31.54	43.19±11.42	40.65±8.49
LF/HF ratio	1.49±2.47	1.07±0.33	1.25±0.50
DFAα_1	1.43±0.28	1.20±0.18	1.17±0.14
DFAα_2	0.70±0.35	0.61±0.30	0.54±0.06
Mean EDA (µS)	0.219±0.171	0.234±0.224	0.191±0.379
Tonic component (µS)	0.153±0.123	0.134±0.182	0.026±0.420
Phasic component (µS)	0.296±0.312	0.305±0.221	0.266±0.189

Data are presented as median ± interquartile range

HRV: Heart rate variability, EDA: Electrodermal activity, HR: Heart rate, SDNN: Standard deviation of normal-to-normal intervals, RMSSD: Root mean square of successive differences, VLF: Very-low frequency, LF: Low frequency, HF: High frequency, DFAα_1: Short-term detrended fluctuation analysis, DFAα_2: Long-term detrended fluctuation analysis

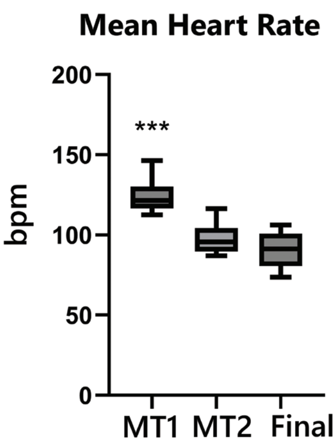

Figure 2. Comparison of mean heart rate across exams. Friedman test followed by Bonferroni-corrected Wilcoxon post-hoc test

****: p<0.001 MT2 and final*

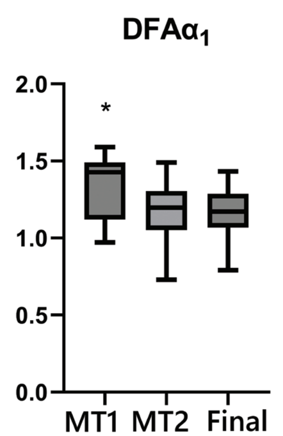

Figure 3. Comparison of DFAα_1 across exams

*Friedman test, *: p < 0.05 MT1 vs. final*

(p=0.0352, Bonferroni-corrected), further supporting the notion of a front-loaded autonomic load.

Non-linear HRV descriptors yielded additional insights into the temporal structure of cardiac variability under academic stress. Both dS1 and TdS, indices reflecting short-term and global reductions in physiological complexity, were markedly elevated during midterm 1 (p=0.0326 and p=0.0211, respectively). Median ± IQR distributions for dS1 were 0.0886±0.1238 (midterm 1), 0.0448±0.0403 (midterm 2), and 0.0792±0.0565 (final); for dS2, 0.0157±0.0406, 0.0188±0.0331, and 0.0243±0.0188; and for TdS, 0.0886±0.1238, 0.0448±0.0403, and 0.0792±0.0565 across the three exam sessions (Figure 4). Collectively, these signatures point to a transient disruption of physiological complexity specifically at the onset of the examination sequence.

Thermoregulatory indices exhibited a comparable pattern of session-related fluctuations. Median ± IQR values for body temperature were 26.58±1.40 °C (midterm 1), 27.50±1.94 °C (midterm 2), and 27.62±3.21 °C (final), indicating modest yet structured variation across assessment periods (Figure 5). These findings underscore the sensitivity of peripheral temperature to examination-induced psychophysiological load and provide an ancillary lens through which to contextualize autonomic reactivity.

Taken together, these results delineate a coherent physiological profile in which early-session examinations elicit heightened sympathetic signatures and reduced autonomic complexity, with subsequent attenuation of these signatures across the exam trajectory. Such temporal dynamics align with theoretical models of anticipatory stress and habituation within real-world cognitive performance contexts.

Relationship Between Exam Scores and Physiological Parameters

Table 2 delineates the associations between exam performance and electrophysiological stress markers derived from HRV and EDA. Correlation analyses demonstrated a statistically significant inverse association between LF power, reflecting baroreflex-mediated sympathetic vasomotor modulation, and exam scores (ρ=-0.440, p=0.015). Likewise, the LF/HF ratio, a canonical index of sympathovagal balance, exhibited a significant negative correlation with performance (ρ=-0.363, p=0.048), indicating that a shift toward sympathetic predominance was linked to reduced academic outcomes.

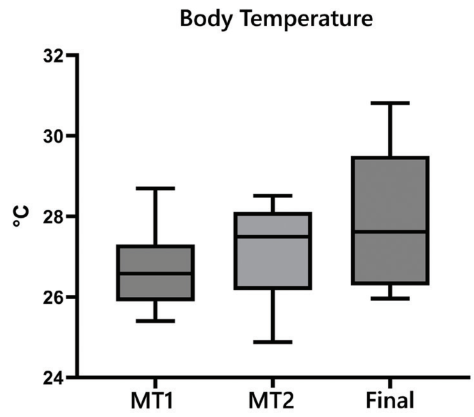

Figure 5. Comparison of body temperature across exams. Values represent wrist skin temperature recorded by the Empatica E4, which may be environmental conditions. Peripheral skin temperature did not differ significantly across examination sessions (Friedman test, p>0.05). Therefore, observed variations should be interpreted cautiously because of potential environmental influences

Table 2. Relationship between exam scores and stress markers

Parameters	Correlation coefficient (ρ)	p-value
LF power	-0.440	0.015
LF/HF ratio	-0.363	0.048
DFAα1	0.41	0.049

LF: Low frequency, HF: High frequency, DFAα1: Short-term detrended fluctuation analysis

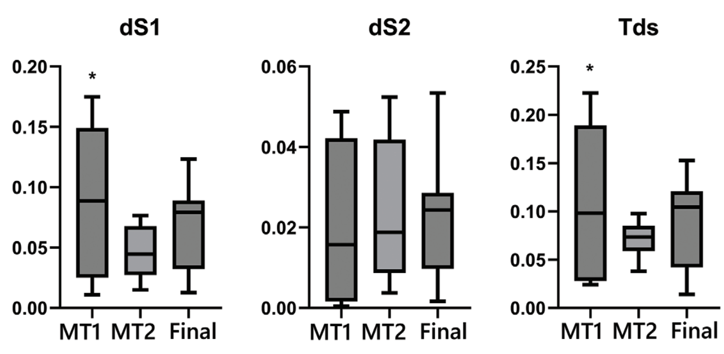

Figure 4. Comparison of physiological complexity loss parameters across exams

dS1: Short-term physiological complexity loss parameter derived from DFAα1, dS2: Long-term physiological complexity loss parameter derived from DFAα2, Tds: Total physiological complexity loss. Statistical analysis was performed using the Friedman test to assess differences across exam sessions (MT1, MT2, final). For parameters reaching statistical significance, post-hoc pairwise comparisons were conducted using the Wilcoxon matched-pairs test with Bonferroni correction. Asterisks indicate statistically significant differences relative to other exam sessions (: p<0.05)*

Conversely, $DFA\alpha_1$, a non-linear metric capturing short-term fractal scaling properties and reflecting the system's intrinsic autonomic complexity, showed a significant positive correlation with exam scores ($\rho=0.410$, $p=0.049$). This pattern suggests that the preservation of fractal-like heartbeat dynamics and reduced autonomic rigidity may confer a performance advantage under evaluative stress.

Discussion

Analysis of electrophysiological data collected from 10 university students during three examination sessions revealed session-dependent variations in autonomic indices. Several HRV measures, including $DFA\alpha_1$ and the HRV-derived stress index, differed across sessions, suggesting that cardiac autonomic dynamics are sensitive to examination context. However, given the exploratory design, a small and homogeneous sample, and the absence of baseline physiological recordings, these findings should be interpreted as descriptive associations rather than evidence of stress-induced physiological effects.

These observations are broadly consistent with prior studies conducted under naturalistic examination conditions. Dimitriev et al.[30] reported reductions in HRV complexity and elevations in $DFA\alpha_1$ during examinations, with $DFA\alpha_1$ associated with state anxiety rather than performance outcomes. Similarly, Kazdagli et al.[3] observed elevated $DFA\alpha_1$ values in students with higher exam-related anxiety, alongside higher exam scores in some individuals, indicating that $DFA\alpha_1$ does not uniformly reflect maladaptive autonomic dysregulation but may coexist with effective cognitive engagement under evaluative stress. Thus, $DFA\alpha_1$ appears to function as a stress-responsive marker whose relationship with performance is context-dependent and moderated by individual stress-regulation capacity.

Consistent with this interpretation, $DFA\alpha_1$ showed a positive association with exam scores, whereas LF power and the LF/HF ratio were negatively correlated with performance, linking sympathetic dominance to poorer academic outcomes[3]. This pattern aligns with a non-linear arousa-performance relationship, whereby moderate autonomic activation may support cognitive efficiency while excessive sympathetic dominance impairs it, in line with the Yerkes-Dodson framework[31]. As emphasized in previous reviews, such physiological-performance associations are typically modest and influenced by interindividual variability[32].

Among non-linear HRV metrics, TdS was elevated during the initial examination, suggesting greater disruption of autonomic complexity during early evaluative stress, consistent with prior reports of physiological complexity loss under stress[29,33]. The subsequent reduction in TdS across later sessions may reflect partial habituation; however, interpretation is constrained by the lack of baseline recordings and the use of wrist-based PPG-derived HRV signals, which are susceptible to motion artifacts.

EDA findings further supported sensitivity to exam-related arousal. As a peripheral index of sympathetic activation, EDA reflects emotional engagement and cognitive load[19,24] and higher tonic EDA levels have been reported in students with elevated test anxiety[3]. Nevertheless, EDA represents a nonspecific arousal signal influenced by environmental and individual factors[34,35], exhibits delayed responses relative to sympathetic activation[36], and lacks full methodological standardization across studies[37]. Accordingly, its interpretive value is greatest when integrated with complementary autonomic measures such as HRV[38-41].

Skin temperature provided additional, context-dependent information. Reductions observed in some participants are consistent with sympathetic vasoconstriction; however, substantial interindividual variability and sensitivity to environmental conditions limit its utility as a standalone marker of autonomic arousal[33].

Overall, this study provides an integrative, exploratory assessment of exam-related physiological patterns using wearable-derived measures of HRV, EDA, and skin temperature. While session-dependent autonomic variations and associations with academic performance were observed, conclusions are limited by sample size, the absence of baseline and subjective stress measures, environmental confounding, and reliance on offline signal processing. Future studies incorporating larger and more diverse cohorts, baseline-controlled and longitudinal designs, multimodal psychological–physiological assessments, and advanced analytic approaches, including artificial intelligence-based methods, will be essential to refine the use of wearable technologies for monitoring and managing exam-related stress.

Study Limitations

Several limitations of the present study should be acknowledged. First, the study relies on a secondary analysis of a publicly available dataset with a relatively small sample size ($n=10$) consisting of a single student cohort, which constrains the statistical power and the generalizability of the findings across broader and more diverse student populations. Second, the dataset lacked non-exam resting baseline physiological recordings; therefore, all autonomic evaluations were restricted to within-session and between-session comparisons rather than deviations from true resting autonomic states. Third, concurrent psychological assessments (such as validated self-report scales for state anxiety, test anxiety, or subjective stress perception) were not available, limiting the direct triangulation between subjective psychological experiences and physiological markers. Fourth, physiological data were acquired exclusively through wrist-worn wearable PPG and dry electrodermal electrodes (Empatica E4). Although ecologically valid and non-invasive, wrist PPG is inherently more susceptible to motion artifacts and subtle peak-detection errors than standard ECG, and peripheral skin conductance and temperature are subject to uncontrolled ambient and environmental factors. Finally, the current analyses were performed post-hoc and offline; real-time stress detection and continuous longitudinal monitoring across multiple academic periods remain to be implemented in future research.

Conclusion

In conclusion, this study demonstrates that academic examinations elicit dynamic, session-dependent autonomic nervous system responses that can be non-invasively tracked using wearable multi-sensor technology. Our findings indicate that heightened sympathetic dominance (reflected by elevated LF power and LF/HF ratio) is negatively associated with academic performance, whereas the preservation of cardiac autonomic complexity and short-term fractal dynamics (DFAα1) is positively correlated with exam achievement. Furthermore, the temporal trajectory of physiological metrics, characterized by peak cardiovascular arousal and physiological complexity loss during early exam sessions followed by subsequent attenuation, supports models of anticipatory stress and subsequent adaptation in evaluative academic environments. Combining HRV, EDA components, and peripheral temperature provides a multidimensional psychophysiological profile that surpasses single-parameter assessments. Despite the exploratory nature of this work, these results highlight the potential of wearable biosensors as objective tools for educational psychophysiology. Future research employing larger cohorts, controlled baseline protocols, multimodal psychological-physiological tracking, and real-time artificial intelligence-assisted analytics will be instrumental in developing individualized stress-monitoring frameworks and adaptive pedagogical interventions to enhance student well-being and academic performance.

Ethics

Ethics Committee Approval: Our study was conducted with the approval of the İzmir University of Economics Health Sciences Research Ethics Committee (approval no: E-97429853-050.04-111205, date: 05.11.2025).

Informed Consent: The dataset is available to researchers via the PhysioNet platform, and because participant information is anonymized, an informed consent form is not required.

Footnotes

Authorship Contributions

Design: H.K., Data Collection or Processing: H.K., H.F.Ö., Analysis or Interpretation: H.K., H.F.Ö., Literature Search: H.K., H.F.Ö., Writing: H.K., H.F.Ö.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

References

1. Crego A, Carrillo-Diaz M, Armfield JM, Romero M. Stress and academic performance in dental students: the role of coping strategies and examination-related self-efficacy. J Dent Educ. 2016;80:165-72.
2. Beltrán-Velasco AI, Bellido-Esteban A, Ruisoto-Palomera P, Clemente-Suárez VJ. Use of portable digital devices to analyze autonomic stress response in psychology objective structured clinical examination. J Med Syst. 2018;42:35.
3. Kazdagli H, Ozel HF, Ozturk S, et al. Electrophysiological detection of exam stress in health schools' students. Physiol Int. 2024;111:222-41.
4. Akanpaadgi E, Binpimbu F, Kuuyelleh EN. The impact of stress on students' academic performance. Eureka J Educ Res. 2023;2:60-7.
5. Lee SM, Oh Y. The mediator role of perceived stress in the relationship between academic stress and depressive symptoms among e-learning students with visual impairments. Journal of Visual Impairment & Blindness. 2017;111:123-34.
6. Shrivastava U, Thakur A, Singh TP, et al. Comparative analysis of stress levels among undergraduate students at St. John's College, Agra. Cureus. 2024;16:e67452.

7. Rosaline MD, Anggraeni DT. Factors related to academic stress during the covid-19 pandemic in nursing students of UPN Veteran Jakarta. Advances in Health Sciences Research.2020;30:374-9.
8. Joseph N, Nallapati A, Machado MX, et al. Assessment of academic stress and its coping mechanisms among medical undergraduate students in a large Midwestern University. Curr Psychol. 2021;40:2599-609.
9. Nagle YK, Sharma U. Academic stress and coping mechanism among students: an Indian perspective. J Child Adolesc Psychol. 2018;2:6-8.
10. Bedewy D, Gabriel A. Examining perceptions of academic stress and its sources among university students: the perception of academic stress scale. Health Psychol Open. 2015;2:2055102915596714.
11. Razavi T. Self-report measures: an overview of concerns and limitations of questionnaire use in occupational stress research. 2001:1-29.
12. Laferton JA, Stenzel NM, Fischer S. The Beliefs about stress scale (BASS): development, reliability, and validity. Int J Stress Manag. 2018;25:72.
13. Zunhammer M, Eberle H, Eichhammer P, Busch V. Somatic symptoms evoked by exam stress in university students: the role of alexithymia, neuroticism, anxiety and depression. PLoS One. 2013;8:e84911.
14. Danthony S, Mascret N, Cury F. Development and validation of a scale assessing test anxiety in physical education. J Teach Phys Educ. 2019;38:357-66.
15. Murphy L, Denis R, Ward CP, Tartar JL. Academic stress differentially influences perceived stress, salivary cortisol, and immunoglobulin-a in undergraduate students. Stress. 2010;13:366-71.
16. Guimarães T, Silva AP, Monteiro I, Gomes R. Stress in university students and cardiovascular response to academic stressors. 2014.
17. Obi A, Nwobodo EO, Dimkpa U, Maduka SO, Fintan E. Assessment of stress level of young undergraduates before and after a degree examination using heart rate variability analysis. European Journal of Medical and Health Sciences. 2021;3:1-6.
18. Hammoud S, Karam R, Mourad R, Saad I, Kurdi M. Stress and heart rate variability during university final examination among lebanese students. Behavioral Sciences. 2019;9:3.
19. Pop-Jordanova N, Pop-Jordanov J. Electrodermal activity and stress assessment. Pril (Makedon Akad Nauk Umet Odd Med Nauki). 2020;41:5-15.
20. Sikder MSBB, Salimullah SM, Ahmed I, Hossain D, Rajib MA, Wadud K. Galvanic skin response based cognitive load measurement for online exams. IEEE. 2021.
21. Horvers A, Tombeng N, Bosse T, Lazonder AW, Molenaar I. Detecting emotions through electrodermal activity in learning contexts: a systematic review. Sensors (Basel). 2021;21:7869.
22. Al Jammali Z, Al Yasiry A, Alsaadi MA, Alyassery AAA. The impact of anxiety on undergraduate students' outcomes during prosthetic exam. J Stomatol. 2024;77:29-33.
23. Roos AL, Goetz T, Krannich M, et al. Control, anxiety and test performance: self-reported and physiological indicators of anxiety as mediators. Br J Educ Psychol. 2023;93:72-89.
24. Rahma ON, Putra AP, Rahmatillah A, et al. Electrodermal activity for measuring cognitive and emotional stress level. J Med Signals Sens. 2022;12:155-62.
25. Amin R, Wickramasuriya D, Faghih RT. A wearable exam stress dataset for predicting cognitive performance in real-world settings (version 1.0.0). PhysioNet. 2022.
26. Amin R, Wickramasuriya DS, Faghih RT. A wearable exam stress dataset for predicting grades using physiological signals. IEEE. 2022.
27. Shaffer F, Ginsberg JP. An overview of heart rate variability metrics and norms. Front Public Health. 2017;5:258.
28. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Circulation. 1996;17:354-81.
29. Ozel HF, Kazdagli H. A simple approach to determine loss of physiological complexity in heart rate series. Biomed Phys Eng Express. 2023;9.
30. Dimitriev DA, Saperova EV, Dimitriev AD. State anxiety and nonlinear dynamics of heart rate variability in students. PLoS One. 2016;11:e0146131.
31. Corbett M. From law to folklore: work stress and the Yerkes-Dodson Law. J Manag Psychol. 2015;30:741-52.
32. Kim HJ, Park Y, Lee J. The validity of heart rate variability (HRV) in educational research and a synthesis of recommendations. Educ Psychol Rev. 2024;36:42.
33. Marazziti D, Di Muro A, Castrogiovanni P. Psychological stress and body temperature changes in humans. Physiol Behav. 1992;52:393-5.
34. Colbert AP, Spaulding K, Larsen A, Ahn AC, Cutro JA. Electrodermal activity at acupoints: literature review and recommendations for reporting clinical trials. J Acupunct Meridian Stud. 2011;4:5-13.
35. Li S, Sung B, Lin Y, Mitas O. Electrodermal activity measure: a methodological review. Ann. Tour. Res. 2022;96:103460.
36. Benedek M, Kaernbach C. A continuous measure of phasic electrodermal activity. J Neurosci Methods**.** 2010;190:80-91.
37. Tronstad C, Amini M, Bach DR, Martinsen ØG. Current trends and opportunities in the methodology of electrodermal activity measurement. Physiol Meas. 2022;43.
38. Cecchi S, Piersanti A, Poli A, Spinsante S. Physical stimuli and emotions: EDA features analysis from a wrist-worn measurement sensor. IEEE. 2020.
39. Lee J, Hwang HB, Lee S, et al. Analysis of acute stress reactivity and recovery in autonomic nervous system considering individual characteristics of stress using HRV and EDA. IEEE Access. 2024.
40. Klimek A, Mannheim I, Schouten G, Wouters EJM, Peeters MWH. Wearables measuring electrodermal activity to assess perceived stress in care: a scoping review. Acta Neuropsychiatr. 2023;37:e19.
41. K Kasap A, Kurt B. Exploring the correlation of physiological stress signals with student exam performance: a preliminary study. Appl Psychophysiol Biofeedback. 2025;50:149-64.