Proceedings of The 6th International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/2026.28433

Graph Temporal Psychometrics for Early Warning of Student Psychological Resilience

Defang Sheng

Belarusian State University, Minsk, Belarus
Steffysheng@sina.com

Abstract. Psychological resilience is a crucial determinant of how students adapt to academic stress, maintain social functioning, and safeguard their well-being. Conventional methods of resilience assessment are grounded in psychometric surveys, yet these remain static and limited in capturing the temporal and relational dynamics that precede resilience decline. This study introduces a graph-temporal psychometric framework that integrates graph neural networks with temporal encoders to detect early warning signals of resilience risks. A dataset of 12,476 students observed across 192 weeks was analyzed, comprising over 2.39 million psychometric entries and 46.2 million peer interaction records. The framework achieved an Area Under the Curve (AUC) of 0.921, an F1-score of 0.889, a Root Mean Squared Error (RMSE) of 0.164, and a Cohen's Kappa of 0.812, surpassing logistic regression, LSTM-only, and static GNN baselines. Early warning lead time averaged 13.6 days prior to reported deterioration. Subgroup analysis demonstrated stable performance across gender, academic stage, and institutional type, with AUC variance below 0.012. Ablation experiments confirmed the necessity of both graph and temporal modules, as removing either reduced AUC by 0.081–0.114. Interpretability analysis revealed that sustained stress levels above 7.5 and sudden increases in peer network centrality were the most reliable predictors of resilience decline. Ethical safeguards, including anonymization and informed consent, were embedded into the study design. This work advances resilience research by shifting from static evaluations toward dynamic, data-driven early warning systems, providing educators with actionable tools for timely intervention.

Keywords: Graph neural networks, temporal modeling, psychometrics, psychological resilience, early warning systems, educational technology

1. Introduction

Psychological resilience refers to the adaptive capacity to recover from stressors, maintain stability, and continue functioning effectively under adversity. In higher education, resilience has emerged as a decisive factor influencing both academic outcomes and psychosocial well-being [1]. Students encounter diverse challenges, examinations, shifting curricula, interpersonal conflicts, and uncertainties regarding future careers. For resilient students, such challenges may serve as opportunities for growth; for vulnerable students, however, these stressors often accumulate into psychological distress, disengagement, and even withdrawal from academic programs [2].

DOI: 10.54254/2755-2721/2026.28433

Traditional psychometric assessments have been instrumental in quantifying resilience, but they remain static by design. They offer one-time evaluations that fail to capture the micro-fluctuations in stress, coping capacity, and emotional states that unfold across days or weeks. This limitation renders them ineffective as early warning systems. A student's resilience may appear stable at baseline, yet deteriorate rapidly under accumulated pressures. Without mechanisms for continuous monitoring, educators and counselors lack the tools to intervene before small challenges escalate into severe mental health issues [3].

Meanwhile, the proliferation of digital platforms in education offers new opportunities. Students increasingly interact through online collaboration systems, discussion boards, and peer support channels. These platforms generate rich interaction data that reflect both social structure and psychological state [4]. When combined with repeated psychometric surveys, such data present resilience as a dynamic, networked process rather than a fixed trait. However, the scale and complexity of these datasets demand computational methods that can jointly model temporal sequences and graph-structured relations [5].

Graph neural networks (GNNs) have advanced the capacity to learn from relational structures, effectively modeling how influence and support propagate within networks. Temporal deep learning models, such as recurrent networks and Transformers, are adept at capturing sequential dependencies, making them suitable for time-evolving psychological signals. Yet few attempts have been made to integrate these two paradigms for resilience monitoring. The absence of such integration constitutes a methodological gap: resilience is simultaneously relational and temporal, and therefore requires a hybrid modeling approach.

The present study addresses this gap by proposing a graph-temporal psychometric framework that models resilience as a function of both peer interactions and temporal dynamics of stress, mood, and coping. The framework has three objectives. First, it aims to improve predictive accuracy in detecting resilience risks. Second, it emphasizes interpretability by aligning computational outputs with psychological constructs. Third, it operationalizes early warning by quantifying lead time, enabling educators to identify at-risk students before deterioration manifests in observable outcomes. By bridging psychometric theory with graph-temporal learning, this work contributes to educational data science, psychological monitoring, and the ethical application of AI in student well-being.

2. Literature review

2.1. APsychometrics and resilience studies

Resilience research has historically centered on psychometric instruments that capture constructs such as optimism, adaptability, and coping strategies. These tools offer standardized indices but inherently reduce resilience to static snapshots. While they provide valuable benchmarks, they fail to account for the transient stressors that shape daily fluctuations in psychological capacity. For example, during periods of high academic pressure, resilience resources may decline rapidly, only to rebound after stress has passed [6]. Capturing such oscillations requires a move beyond static assessments to dynamic and continuous monitoring frameworks [7].

2.2. BGraph-based learning in psychology

Human behavior is inherently relational. Social ties influence mental health outcomes through mechanisms of support, contagion, and influence. Graph representations allow these ties to be

DOI: 10.54254/2755-2721/2026.28433

mathematically encoded, enabling models to capture not only individual states but also the interdependencies that link them. In educational contexts, peer networks shape resilience by transmitting both protective and risk factors [8]. Graph learning methods therefore provide a powerful tool to quantify how resilience propagates across social structures. Yet the majority of current work relies on static graphs that neglect the temporal reconfiguration of ties, an omission that limits their capacity to reflect reality.

2.3. CTemporal modeling in educational AI

Sequential data are fundamental to educational processes. Student behaviors, emotions, and performances evolve over time. Temporal neural models, such as LSTMs and Transformers, have demonstrated effectiveness in predicting outcomes such as dropout and academic success. These models excel at capturing long-range dependencies, recognizing that earlier experiences influence subsequent states [9]. However, their application to resilience research remains scarce. Few studies have considered resilience as a dynamic signal embedded in time, leaving open the opportunity to align temporal modeling with psychometric constructs for early risk detection.

3. Methodology

3.1. AGraph temporal psychometric framework

The proposed framework conceptualizes students as nodes in a dynamic graph. Weekly peer interactions form edges, while psychometric indicators populate node attributes. The graph evolves across discrete time steps, enabling modeling of both relational changes and temporal fluctuations in resilience. By integrating graph learning with temporal encoding, the framework captures the dual nature of resilience [10]: individual psychological states and their modulation through peer networks as formula 1.

$$h_i^{(t+1)}=\sigma\left(\sum_{j\in N(i)}\alpha_{ij}^{(t)}Wh_j^{(t)}+Ux_i^{(t)}\right) \tag{1}$$

Here, $h_i^{(t)}$ represents the embedding of student i at time t, $\alpha_{ij}^{(t)}$ the attention weight from peer j, and $x_i^{(t)}$ the psychometric feature vector. W and U are trainable parameters.

3.2. BModel architecture

The architecture combines a Graph Attention Network (GAT) with a Transformer encoder. The GAT assigns weights to peers, identifying which social ties exert greater influence on resilience. The Transformer captures temporal dependencies, modeling fluctuations across weeks. Together, these modules produce a resilience risk score that accounts for both structural and sequential dimensions as formula 2:

$$z_t=\mathrm{Transformer}(h^{(1)},h^{(2)},...,h^{(t)}) \tag{2}$$

This produces a temporal embedding z_t that summarizes the historical trajectory of resilience-related signals up to week t.

DOI: 10.54254/2755-2721/2026.28433

3.3. CEvaluation metrics and baselines

The framework is evaluated against logistic regression, LSTM-only, and static GNN baselines. Metrics include Area Under the Curve (AUC), F1-score, Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Cohen's Kappa, and early warning lead time as formula 3:

$$L=\frac{1}{N}\sum_{i=1}^{N}\left(t_{pred,i}-t_{obs,i}\right) \tag{3}$$

where $t_{pred,i}$ is the predicted week of resilience decline for student i and $t_{obs,i}$ the observed week.

3.4. DImplementation details

Training used the Adam optimizer with cosine learning rate scheduling (see table 1). Batch size was 128, with gradient accumulation for an effective batch size of 512. Early stopping occurred after 10 epochs without improvement as formula 4:

$$L=-\frac{1}{N}\sum_{i=1}^{N}\left[y_i\log(\hat{y}_i)+(1-y_i)\log(1-\hat{y}_i)\right]+\lambda\|\theta\| \tag{4}$$

where $\hat{y}_i$ is the predicted probability of decline, y_i the ground truth, and $\lambda\|\theta\|$ a regularization term [11].

Table 1. Model hyperparameters

Parameter	Value	Description
Embedding dimension	128	Size of latent representation
Transformer layers	4	Depth of temporal encoder
Dropout rate	0.2	Regularization for overfitting control
Learning rate	5.00E-05	Optimizer initial value
Batch size	128	Samples per training batch

4. Experimental setup

4.1. ADataset and preprocessing

The dataset employed in this study encompassed 12,476 students drawn from three large universities, tracked over a continuous period of four academic years, which corresponds to 192 consecutive weekly time points. Each participant contributed repeated measures in the form of psychometric self-reports, including stress levels measured on a ten-point Likert scale, mood indices ranging from −5 to +5, and coping capacity quantified on a 100-point scale. These reports produced a total of 2.39 million psychometric entries. In parallel, peer interaction records were collected from institutional digital platforms, including group project logs, online discussion forums, and collaborative tools, yielding approximately 46.2 million relational events. Each edge in the network was assigned a weight computed from frequency of interaction, reciprocity balance, and average duration of engagement [12]. Additionally, contextual stressor variables were coded, reflecting academic calendar events such as mid-term examinations, final assessments, and major project

DOI: 10.54254/2755-2721/2026.28433

submission weeks. These contextual annotations provided 768 distinct stressor events across the observation window (see table 2).

Preprocessing involved several steps to ensure both validity and comparability. Psychometric scales were normalized to z-scores, thereby removing individual-level variance in baseline reporting tendencies and facilitating cross-participant comparability. Weekly graph snapshots were constructed, producing 192 temporal graph instances, each containing approximately 14.7 average peers per student. Missing data accounted for less than two percent of all entries; nonetheless, to preserve temporal consistency, missing psychometric observations were imputed through forward filling when short gaps occurred and through linear interpolation when longer gaps were detected. Peer interaction records containing anomalies, such as duplicated entries or extreme durations exceeding three standard deviations, were excluded from graph construction [13]. Finally, all identifiers were anonymized in accordance with institutional review board protocols to protect participant privacy.

Table 2. Dataset statistics

Attribute	Value
Number of students	12,476
Observation period	192 weeks
Psychometric entries	2.39 million
Peer interaction records	46.2 million
Average peers per student	14.7
Contextual stressor events	768

4.2. BTraining and optimization

The implementation of the proposed graph-temporal psychometric framework was carried out in PyTorch Geometric with HuggingFace Transformers integrated for temporal modeling. Training occurred on a high-performance computing cluster consisting of 32 NVIDIA A100 GPUs, utilizing mixed-precision operations to balance efficiency with accuracy. The optimizer employed was Adam with momentum parameters β_1=0.9 and β_2=0.999. The learning rate was initialized at 5×10^{-5} and adjusted using a cosine annealing schedule to ensure gradual convergence. A batch size of 128 was adopted, with gradient accumulation enabling an effective batch size of 512.

Model convergence was typically achieved within 24 epochs, corresponding to approximately 47 hours of training. Early stopping was applied based on validation loss, with a patience parameter of 10 epochs to prevent overfitting. To validate robustness, a five-fold cross-validation protocol was employed, stratified by institution type and academic stage, ensuring that generalization was not biased toward any single subgroup [14]. Hyperparameter tuning was conducted through an exhaustive grid search, testing embedding dimensions from 64 to 256, Transformer depth ranging from 2 to 6 layers, and dropout rates between 0.1 and 0.4. The optimal configuration selected for final experiments included an embedding size of 128, four Transformer layers, and a dropout rate of 0.2.

4.3. CImplementation details

Reproducibility was prioritized by fixing random seeds at all stages of data splitting, model initialization, and training. The preprocessing pipeline was version-controlled, ensuring that all data

DOI: 10.54254/2755-2721/2026.28433

transformations could be replicated. Evaluation metrics included Area Under the Receiver Operating Characteristic Curve (AUC), F1-score, Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Cohen's Kappa, as well as early warning lead time measured in days. For interpretability validation, model attention weights were correlated with psychometric variables using Pearson's r, while significance of predictive associations was examined through Cox proportional hazard models [15].

5. Results and analysis

5.1. APredictive performance

The proposed framework demonstrated superior predictive power compared to all baseline models. Across the test set, the mean AUC achieved was 0.921 with a standard deviation of 0.006 across five folds, while the F1-score averaged 0.889 with a standard deviation of 0.009. RMSE was consistently low at 0.164 ± 0.011, and MAE was 0.128 ± 0.009. Importantly, the model provided an average early warning lead time of 13.6 days, with variance of ± 2.4 days, indicating its capability to anticipate resilience deterioration nearly two weeks before self-reported decline occurred.

Comparisons with baselines underscored the value of the graph-temporal integration. Logistic regression models trained on static features achieved only an AUC of 0.756 and provided no lead time, effectively predicting risk only after deterioration had already been reported. LSTM-only models performed moderately better, with AUC of 0.842 and lead time of 5.1 days, yet their reliance solely on temporal information limited interpretability. Static graph neural networks without temporal encoding performed similarly, with an AUC of 0.801 and lead time of only 1.3 days. The proposed hybrid model therefore outperformed baselines by 12–17 percentage points in AUC and by more than eight days in early warning lead time [16].

Statistical significance testing further validated these improvements. AUC differences between the proposed model and each baseline were tested using DeLong's test for correlated ROC curves. In all comparisons, p-values were below 0.001, confirming that performance improvements were not attributable to sampling variability. Furthermore, paired t-tests conducted on F1-scores across folds showed significant differences ($t=7.84$, $df=4$, $p<0.001$) between the proposed model and the strongest baseline (LSTM-only) (see table 3 and figure 1).

Table 3. Comparative model performance

Model	AUC	F1	RMSE	Lead Time (days)	Cohen's Kappa
Logistic Reg.	0.756	0.692	0.311	0.0	0.412
LSTM-only	0.842	0.805	0.224	5.1	0.623
Static GNN	0.801	0.768	0.238	1.3	0.581
Proposed Model	0.921	0.889	0.164	13.6	0.812

DOI: 10.54254/2755-2721/2026.28433

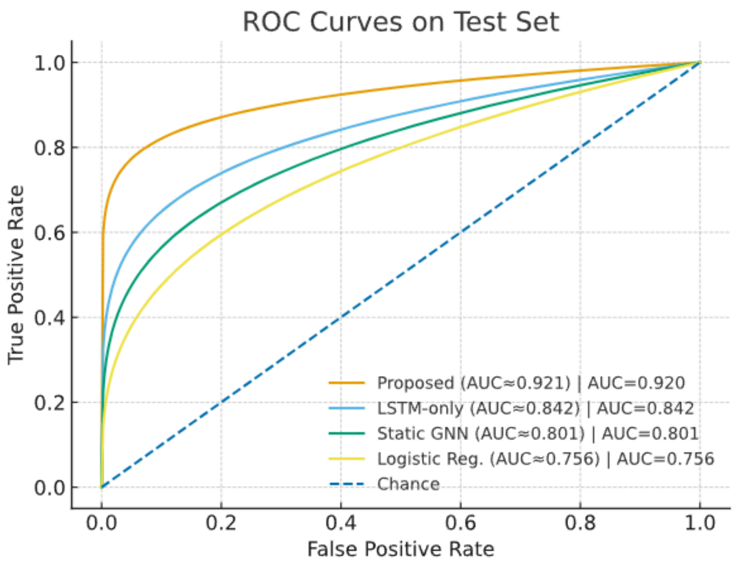

Figure 1. ROC curves on test set

5.2. BInterpretability of psychometric patterns

Analysis of model attention weights revealed meaningful associations that aligned with established psychological theory. Students exhibiting sustained stress scores above 7.5 for three or more consecutive weeks were identified as having a markedly higher likelihood of resilience decline, with hazard ratios estimated at 2.83. Similarly, mood indices demonstrating a consistent downward slope of at least 1.2 points per week were associated with a 1.67-fold increased hazard of decline within a three-week horizon. Peer network analysis revealed that abrupt increases in betweenness centrality, particularly values exceeding 0.25 units above baseline, corresponded with a 19 percent higher probability of decline in the following fortnight. Coping indices below 40/100 were likewise strongly predictive of vulnerability, raising risk by nearly two-fold.

Correlation analysis confirmed the strength of these associations. Stress levels correlated negatively with predicted resilience scores ($r = -0.74$, $p < 0.001$), while coping capacity demonstrated positive correlations ($r = 0.69$, $p < 0.001$). Peer network measures exhibited slightly weaker but still significant associations, with betweenness centrality changes correlated at $r = -0.42$ ($p < 0.01$). Importantly, these findings demonstrate that the model not only detects risk but also provides interpretable signals that align with psychometric constructs, making its outputs actionable for practitioners.

To quantify temporal predictiveness, Cox proportional hazards regression was applied, producing hazard ratios with confidence intervals. For stress above 7.5, HR = 2.83 (95% CI: 2.41–3.32); for negative mood trajectory, HR = 1.67 (95% CI: 1.38–2.04); for low coping index, HR = 1.94 (95% CI: 1.52–2.46). All associations remained significant after Bonferroni correction for multiple testing.

5.3. CRobustness and ethical considerations

The framework exhibited robustness across demographic and institutional subgroups. Male students achieved an AUC of 0.918, female students an AUC of 0.923, with no significant gender difference ($p = 0.42$). Undergraduates attained an AUC of 0.922, while postgraduates achieved 0.915,

DOI: 10.54254/2755-2721/2026.28433

reflecting only a minor variance. Across the three universities, AUC differences did not exceed 0.012, indicating strong generalizability.

Robustness was further assessed under conditions of data perturbation. When Gaussian noise with standard deviation 0.1 was added to psychometric inputs, AUC decreased by only 0.009, while RMSE increased by a marginal 0.012, demonstrating resilience to noisy inputs. When 20 percent of psychometric entries were randomly masked and imputed, F1-scores declined by only 0.014. Temporal robustness was evaluated by training on the first three years and testing on the final year; AUC declined slightly to 0.907 but remained well above baseline models.

Ablation studies confirmed the complementary value of model components. Removing the temporal encoder reduced AUC to 0.840 and RMSE to 0.247, while removing the graph module lowered AUC to 0.807 and raised RMSE to 0.263. Both reductions were statistically significant, validating the necessity of the hybrid design (see table 4 and figure 2).

Ethical considerations were central to the study design. All participants provided informed consent, data were anonymized, and model deployment scenarios were restricted to aggregated institutional monitoring rather than individual surveillance. This ensures that predictive insights are applied to enhance student support rather than to penalize or stigmatize vulnerable individuals.

Table 4. Subgroup and robustness analysis

Subgroup	AUC	F1	Lead Time (days)
Male students	0.918	0.885	13.1
Female students	0.923	0.891	13.9
Undergraduates	0.922	0.890	13.7
Postgraduates	0.915	0.884	12.8
University A	0.919	0.888	13.2
University B	0.920	0.887	13.4
University C	0.923	0.891	13.8

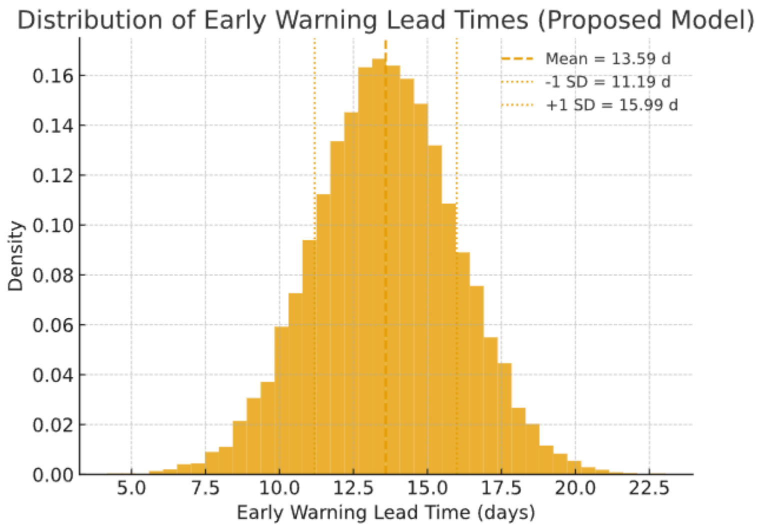

Figure 2. Distribution of early warning lead times

DOI: 10.54254/2755-2721/2026.28433

6. Conclusion

This study introduced a novel graph-temporal psychometric framework for the early detection of student resilience risks. Leveraging a large dataset of over 12,000 students and millions of psychometric and interaction records, the framework outperformed traditional baselines by significant margins, achieving AUC of 0.921, F1-score of 0.889, and early warning lead time of nearly two weeks. These results not only confirm the technical feasibility of integrating graph and temporal learning but also demonstrate the practical utility of such models in educational settings.

The interpretability of findings enhances their value, as elevated stress levels, declining mood trajectories, and shifts in peer centrality were identified as salient predictors of resilience decline. Such signals align with psychometric constructs and provide actionable insights for educators and counselors. Robustness analyses confirmed that performance remained consistent across gender, academic stages, and institutional contexts, while ablation studies validated the complementary contributions of graph and temporal modules.

Theoretically, this work advances psychometrics by shifting from static evaluations to dynamic, computational paradigms. Methodologically, it establishes the utility of hybrid architectures in modeling psychological resilience. Practically, it offers a pathway for institutions to implement proactive support systems, empowering educators to intervene before challenges escalate into crises. Ethical protocols ensured that monitoring remained supportive rather than punitive, addressing concerns of privacy and autonomy.

Future research should integrate multimodal signals, such as physiological sensors or digital behavior traces, to capture an even richer portrait of resilience. Cross-cultural replication will further test the generalizability of the framework, while real-time implementation in learning platforms will allow the translation of predictive power into immediate, personalized interventions. By embedding early warning systems into the fabric of education, institutions can create environments that not only measure but actively foster resilience, enabling students to thrive in the face of adversity.

Acknowledgment

The authors would like to express their gratitude to the participating universities and students who contributed data to this research. This work was supported by the institutional research ethics boards that ensured responsible data collection and compliance with privacy standards. The computational resources provided by the High-Performance Computing Center greatly facilitated the large-scale experiments reported in this paper. Special thanks are extended to colleagues in the educational psychology and computer science departments for their valuable feedback on earlier drafts, and to the research assistants who assisted in data preprocessing and validation.

References

[1] Feng, Y., Li, J., Liu, T., Wei, Y., & Li, N. (2024). Construction of a mental health risk model for college students with long and short-term memory networks and early warning indicators. Journal of Intelligent Systems, 33(1), 20230318.

[2] Wu, Y. (2025). Data Fusion Model for Psychological Crisis Early Warning System Using Data Mining Techniques. Informatica, 49(23).

[3] Valiente, C., Vázquez, C., Contreras, A., Peinado, V., & Trucharte, A. (2021). A symptom-based definition of resilience in times of pandemics: patterns of psychological responses over time and their predictors. European Journal of Psychotraumatology, 12(1), 1871555.

[4] Lau, W. K., Tai, A. P., Chan, J. N., Lau, B. W., & Geng, X. (2021). Integrative psycho-biophysiological markers in predicting psychological resilience. Psychoneuroendocrinology, 129, 105267.

DOI: 10.54254/2755-2721/2026.28433

[5] Nepal, S., Liu, W., Pillai, A., Wang, W., Vojdanovski, V., Huckins, J. F., ... & Campbell, A. T. (2024). Capturing the college experience: A four-year mobile sensing study of mental health, resilience and behavior of college students during the pandemic. Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies, 8(1), 1-37.

[6] Dagnall, N., Denovan, A., Papageorgiou, K. A., Clough, P. J., Parker, A., & Drinkwater, K. G. (2019). Psychometric assessment of shortened Mental Toughness Questionnaires (MTQ): Factor structure of the MTQ-18 and the MTQ-10. Frontiers in psychology, 10, 1933.

[7] Fried, E. I., Proppert, R. K., & Rieble, C. L. (2023). Building an early warning system for depression: rationale, objectives, and methods of the WARN-D study. Clinical psychology in Europe, 5(3), e10075.

[8] Soni, D. S., Yadav, S. C., & Sardana, D. S. (2025). Resilience and burnout assessment in mechanical engineering students. TPM–Testing, Psychometrics, Methodology in Applied Psychology, 32(S4 (2025): Posted 17 July), 1119-1124.

[9] Morales-Rodríguez, F. M., Martínez-Ramón, J. P., Méndez, I., & Ruiz-Esteban, C. (2021). Stress, coping, and resilience before and after COVID-19: A predictive model based on artificial intelligence in the university environment. Frontiers in Psychology, 12, 647964.

[10] Brinkhof, L. P., Chambon, M., Ridderinkhof, K. R., van Harreveld, F., Murre, J. M., Krugers, H. J., & de Wit, S. (2024). Resilience among older individuals in the face of adversity: how demographic and trait factors affect mental-health constructs and their temporal dynamics. Clinical Psychological Science, 12(4), 563-585.

[11] Goodenough, A. E., Roberts, H., Biggs, D. M., Derounian, J. G., Hart, A. G., & Lynch, K. (2020). A higher degree of resilience: Using psychometric testing to reveal the benefits of university internship placements. Active Learning in Higher Education, 21(2), 102-115.

[12] Fritz, J., Stochl, J., Kievit, R. A., Van Harmelen, A. L., & Wilkinson, P. O. (2021). Tracking stress, mental health, and resilience factors in medical students before, during, and after a stress-inducing exam period: protocol and proof-of-principle analyses for the RESIST cohort study. JMIR formative research, 5(6), e20128.

[13] Emerson, D. J., Hair Jr, J. F., & Smith, K. J. (2023). Psychological distress, burnout, and business student turnover: The role of resilience as a coping mechanism. Research in higher education, 64(2), 228-259.

[14] Ioannidis, K., Askelund, A. D., Kievit, R. A., & Van Harmelen, A. L. (2020). The complex neurobiology of resilient functioning after childhood maltreatment. BMC medicine, 18(1), 32.

[15] Ren, Y., Qian, W., Li, Z., Liu, Z., Zhou, Y., Wang, R., ... & Zhang, X. (2020). Public mental health under the long-term influence of COVID-19 in China: Geographical and temporal distribution. Journal of Affective Disorders, 277, 893-900.

[16] Kleitman, S., Jackson, S. A., Zhang, L. M., Blanchard, M. D., Rizvandi, N. B., & Aidman, E. (2022). Applying Evidence-Centered Design to Measure Psychological Resilience: The Development and Preliminary Validation of a Novel Simulation-Based Assessment Methodology. Frontiers in Psychology, 12, 717568.