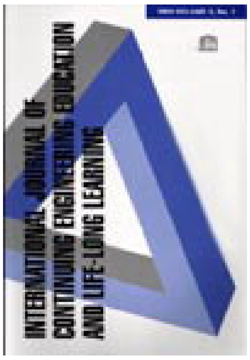


International Journal of Continuing Engineering Education and Life-Long Learning

ISSN online: 1741-5055 - ISSN print: 1560-4624
https://www.inderscience.com/ijceell

How intelligent semi-supervised learning illuminates influencing factors in college students' employment psychology

Yi Zhou

DOI: 10.1504/IJCEELL.2025.10074083

Article History:

Received:	24 February 2025
Last revised:	11 June 2025
Accepted:	09 September 2025
Published online:	28 November 2025

How intelligent semi-supervised learning illuminates influencing factors in college students' employment psychology

Yi Zhou

Applied Technology College,
Soochow University,
Kunshan, 215325, China
Email: zhouyi2017@suda.edu.cn

Abstract: This work proposed an analysis model of the influencing factors of college students' employment psychology combined with intelligent semi-supervised learning technology. The analysis effect of the influencing factors of college students' employment psychology is further improved, helping college students correct their mindset and better cope with social employment. In addition, it introduced the class-aware contrastive learning module and the label-guided iterative self-incremental learning module, which help the model fully explore the potential features of unlabelled data and effectively solve the problem of insufficient labelled data on the psychological factors of college students' employment. It indicated that the higher the mental health literacy of graduates, the higher their psychological resilience level. Therefore, when providing employment guidance, schools need to carry out the work in stages and groups, cultivate students' psychological resilience and positive coping styles in the face of setbacks, and enhance graduates' confidence in their own career development.

Keywords: semi-supervised learning; college students; employment psychology; influencing factors.

Reference to this paper should be made as follows: Zhou, Y. (2025) 'How intelligent semi-supervised learning illuminates influencing factors in college students' employment psychology', *Int. J. Continuing Engineering Education and Life-Long Learning*, Vol. 35, No. 9, pp.52–71.

Biographical notes: Yi Zhou graduated with a Master's degree from Soochow University in June 2017, specialising in Design Studies. She joined the Applied Technology College of Soochow University in September 2017 as an Assistant Professor. Over the past eight years, she has been dedicated to educational research and has led or participated in several provincial and municipal research projects in related fields and has also published over ten academic papers.

1 Introduction

Employment is the biggest livelihood issue and an eternal issue for the development of the times. The continuous expansion of graduates' scale will inevitably affect the quality of talent training and the utilisation of educational resources, and the employment situation of college students will become more and more severe. The research shows that employment pressure has become the biggest psychological pressure that contemporary college students need to face. When many college students are approaching graduation, they have no clear plan for their careers, their employability has not reached the expected level, and their employment pressure is gradually increasing. With the rapid development of economy and the adjustment of industrial structure, industrial enterprises put forward more requirements for high-quality skilled talents. Therefore, college students with low professional ability and comprehensive quality are easily eliminated under the fierce employment situation, and college students' employment is facing severe challenges.

The formulation of employment goal is the first step for college students to start their personal employment choice, and a reasonable employment goal is the foundation of college students' employment and the key to successful employment. Only when college students fully recognise the actual situation of personal special skills, personality characteristics, hobbies, etc., and make a correct personal orientation, can they clarify their advantages and disadvantages in employment choices. At the same time, they need to combine the overall situation of social employment at the present stage and the needs of enterprise development, adopt appropriate job-hunting methods, and form a scientific and objective understanding of the current situation of personal employment, so as to make accurate career choices and future plans (Belle et al., 2022). In addition, in the process of job-hunting practice, when it is found that there is a big gap between personal job-hunting goals and practice, a healthy employment psychology will enable college students to improve their ability to resist stress and frustration, and strive to achieve their job-hunting goals with a more positive attitude and more targeted actions.

At different time nodes and stages, everyone plays different roles, and college students who are about to enter the society should also accept this change. From students to job seekers, college students need to change their personal mentality while adjusting their personal behaviours (Ma et al., 2021). In the process of job hunting, college students will encounter a more colourful social environment than the school, more changeable people to get along with, and more diverse skill requirements. If they still keep the mentality of doing things as students when they look at the people, things and things around them, they will be at a loss in the face of various new situations (Baluku et al., 2021). Therefore, having a healthy employment psychology can help college students quickly adjust their personal mentality and improve their coping ability.

Employment is a social problem that college students must face. With the increasing number of students enrolled in colleges and universities, the scale of running schools continues to expand, higher education presents the trend of mass development, the employment situation of college students is becoming more and more severe, and the psychological pressure of college students' employment is also increasing. Under the new employment situation, college students' employment psychological problems are constantly prominent. As socialist successors, they must strengthen their ability to resist pressure and psychological quality. Based on this, college students should constantly improve their ideological cognition, establish a correct employment concept, reduce their job expectations, and deal with setbacks and challenges in their work with a positive and optimistic attitude. Moreover, colleges and universities should also pay more attention to college students' employment mental health education, establish a perfect employment security system, do a good job in college students' psychological counselling and regulation, help students master professional knowledge and skills, and stimulate their self-confidence. At the same time, the society should also strengthen the care for the employment of college students, and relevant enterprises should implement employment policies, so as to alleviate the employment pressure of college students and help them achieve employment.

This paper proposes an analysis model of college students' employment psychology influencing factors combined with intelligent semi-supervised learning technology, so as to further improve the analysis effect of college students' employment psychology influencing factors through this model and help college students correct their mentality and better cope with social employment. Moreover, this paper introduces class-aware contrastive learning Cacm module and label-guided iterative self-incremental learning Lgis module, which helps the model fully mine the potential characteristics of unlabelled data and effectively solves the problem of insufficient labelled data of college students' employment psychological factors.

2 Related works

2.1 Influencing factors of college students' employment psychology

The reasons for the difficulty of college students' employment are as follows. First, colleges and universities have expanded enrolment and the number of graduates has increased sharply, but the number of jobs is limited. Due to the impact of domestic economic situation, industrial development and other factors, the demand for jobs has decreased significantly, forming a situation of supply less than demand, which is far from meeting the employment needs of college students. Second, the employment market is not standardised and stable enough, and policies such as 'recommendation of excellent students' and 'selection of excellent students' have not been implemented. These adverse factors will seriously affect the fair competition of college students' employment. Third, with the rapid development of society, the requirements for college students are getting higher and higher. The fierce competition in the workplace and complex interpersonal relationships have caused great employment pressure on college students (Alshurideh et al., 2022). Fourth, although many policies and regulations on college students' entrepreneurship have been issued, the policies and regulations do not have a unified management. There are many departments involved in government affairs, the interpretation of the policies is wrong, the implementation of the policies is not enough, and the policy resources are not utilised, which affects the enthusiasm of college students' employment and entrepreneurship. These factors will cause college students' employment mentality imbalance, anxiety, impatience, paranoia and depression.

Colleges and universities pay more attention to the cultivation of students' professional ability, but ignore the cultivation of career planning ability, employment ability, innovation and entrepreneurship ability. First of all, the employment guidance courses offered in freshmen tend to focus on policy education and theoretical education, but lack career planning, career counselling and employment services that students need, or such services lack substantive guidance and operability. Secondly, the talent training program is backward, the teaching method is single, and the emphasis is on theory teaching rather than practical operation ability training. Students lack the awareness and ability to build social capital. Even if the corresponding internship courses are opened in sophomores and juniors, most of them are mere formality, not strong in professionalism, short in time, and students' harvest is not big. The opportunities for college students to have in-depth contact with enterprises are very limited, and often only when they are close to graduation can they learn about enterprises through recruitment seminars. Finally, there is no timely analysis and guidance on the ideological and psychological problems of college students' job selection, and there is no specific and systematic guidance on the employment concept and job selection skills. The integration of ideological and political education and employment education is not close, and there is a 'two skin' phenomenon. Moreover, colleges and universities lack a unified information integration platform and internship platform, there are deficiencies in employment information data management, insufficient publicity and sharing of employment information, and students' access to employment reference information is limited. It can be seen that the employment guidance work can not keep up with the development and changes of students' employment psychology in time, some college students will be indomitable and depressed, some will have the psychology of escaping, lying flat and obeying fate, and the extreme will have mental disorders such as anxiety and depression (Peltz et al., 2021).

The 'slow employment' of students mainly refers to the phased social phenomenon that college students are not in a hurry to find a job after graduation, and choose to review for the postgraduate entrance examination, prepare for the civil service examination, stagger peak employment or take a rest for a period of time before making plans, which is mainly manifested in the temporary unemployment of graduates or the decline in the participation rate of job hunting. College graduates' slow employment, lazy employment and blind employment are also common (Phungsoonthorn and Charoensukmongkol, 2022). First, job seekers' knowledge, ability, psychology and experience are not fully prepared. They are nervous and afraid in the process of job hunting. They dare not recommend themselves boldly and dare not show their ability and personality on their own initiative, which makes it difficult for them to stand out in the competition and suffer repeated setbacks in job hunting. Second, they do not pay attention to positive employment education and employment information at all. They do not listen to the correct information and suggestions from counsellors, school employment guidance service centres, students and employers. Some of them are even arrogant and determined to apply for jobs according to their own ideas (Moriarty et al., 2021). Third, affected by the employment pressure, he signed the contract hastily without knowing the employer. Once he found that things went against his wishes, he regretted. Fourth, they lack a clear understanding of their own interests, abilities and career development direction, which leads them to hesitate and wait and see when choosing a job. Fifth, college students with superior conditions have high self-esteem and ignore the mismatch between their own ability and market demand, resulting in difficulty in finding a suitable job. After graduation, facing the severe employment situation, they are full of worries and do not know what to do, resulting in psychological problems such as fear, impatience, inferiority complex and pride (Kee, 2021).

2.2 Semi supervised learning technology

Semi supervised learning is a branch of machine learning. It is a technology combining fully supervised learning and unsupervised learning. Different calculation methods are used for different data categories. Fully supervised learning is carried out for the marked part of the input data, and unsupervised learning is carried out for the unmarked part (Ryu and Fan, 2023). Literature (Apriceno et al., 2021) divides semi supervised learning algorithms into two categories: inductive learning and transductive learning. Inductive learning, like supervised learning, generates a classification model that can be used to predict sample labels that have not been seen before. Transductive learning does not generate such a model, but directly predicts new samples.

In inductive learning, semi supervised learning method based on disturbance is a hot research direction with high accuracy. This method is based on the semi supervised learning smoothing hypothesis: the labels of other sample points in the small neighbourhood of a sample point should be consistent with the sample point (Liu et al., 2023). Disturbance-based methods are usually implemented using neural networks. They can directly add additional unsupervised loss terms to the objective function, so they can be relatively easily extended to semi supervised learning. LaCosse et al. (2021) proposed a ladder network, adding the unsupervised loss to the loss term of the model. In this model, the feedforward part of the network is used as the encoder for supervised learning, and a decoder is added to reverse the mapping on the encoder. This part supports unsupervised learning. Finally, a new loss term is added to the cost function to punish the reconstruction cost. Zapata-Cuervo et al. (2023) proposed the simple variants of this method, Π-model and temporalassembling. Π-model uses the randomness of dropout to disturb the neural network model, so that the same input will produce different results after passing through the network. Based on the multiple disturbances of the network model, temporal assembling method compares the current iteration prediction value with the exponential moving average of the previous historical iteration prediction value. Easterbrook and Hadden (2021) further improved the temporal assembling method and proposed mean teacher method, which considers the moving average of the model connection weight rather than the moving average of the prediction results. Virtual adversarial training (VAT) (Kotera et al., 2022) proposed a regularisation method considering the direction of sample disturbance. For each data point, whether labelled or not, they will disturb it, making these samples generate anti noise in the input space to realise the disturbance to the samples.

In transductive learning, graph based semi supervised learning is one of the more advanced research directions. The graph-based semi-supervised learning method generally includes three steps: graph creation, graph weighting and reasoning labels (Mbous et al., 2024). Prasath et al. (2021) proposed a Bayesian graph convolution neural network framework, which uses Bayesian method to infer the joint posterior probability of random graph parameters and node labels as the target. Quintiliani et al. (2022) proposed a new framework called graph stochastic neural network, which aims to model the uncertainty of classification functions by simultaneously learning a series of functions (random functions).

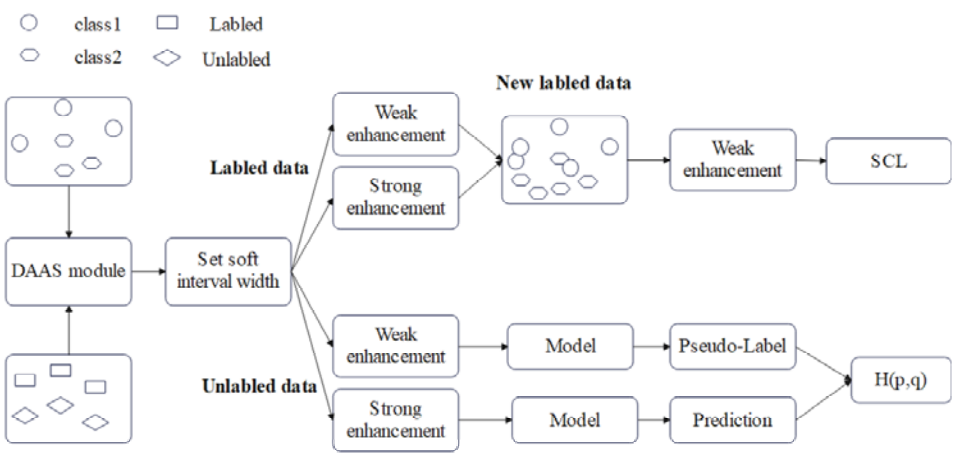
In the existing semi supervised learning methods, the experimental verification part is mostly completed by using image data and graph structure data. These methods do not verify the effectiveness of input for other data types. In addition, although some sample perturbation methods based on data enhancement have achieved good results on image data, these data enhancement methods are only applicable to image data and cannot be extended to non-image data, such as neuropsychological test data (Mascherini et al., 2021).

3 Semi-supervised learning model construction and experiment

3.1 TextMatch model

The framework diagram of TextMatch is shown in Figure 1. The components are mainly divided into the processing of labelled employment psychological data, the processing of unlabelled employment psychological data, the way of enhancing employment psychological data and DAAS module. Next, the content of each component will be introduced in detail.

Figure 1 TextMatch framework diagram (see online version for colours)



For labelled employment psychology data, dynamic employment psychology data enhancement is carried out by DAAS module firstly, and then weak enhancement and strong enhancement are carried out by setting soft interval width, which expands the original labelled employment psychology data. Then, a text instance is selected as an anchor, and then other instances with the same label as it are selected as positive samples from the labelled employment psychology dataset, and instances with different labels as negative samples are selected for supervised comparative learning.

Supcon of the supervised contrastive learning SCL net will be used as the loss function. By comparing the similarity and difference between different employment psychological data samples, richer and more robust feature expressions can be effectively extracted. The loss function L_S is shown in formulas (1)–(2) (Charoensukmongkol and Phungsoonthorn, 2022).

$$L_S = \sum_{i=1}^{2N} L_i^{sup} \tag{1}$$

$$L_S = \frac{1}{2N_{\tilde{y}_i} - 1} \sum_{j=1}^{2N} 1_{i \neq j} \cdot 1_{\tilde{y}_i} \tilde{y}_j \cdot \log \frac{\exp(z_i \cdot z_j / \tau)}{\sum_{j=1}^{2N} 1_{i \neq j} \cdot \exp(z_i \cdot z_j / \tau)} \tag{2}$$

For unlabelled employment psychological data, dynamic employment psychological data enhancement is also carried out by DAAS module, and the width of soft interval is set to distinguish between strong and weak enhancement. For weakly enhanced samples, the threshold with high confidence is screened by the model as a false label and retained. Strong enhancement will be performed on the same employment psychological data, and cross-entropy loss will be performed on the strongly enhanced prediction and the weakly enhanced pseudo-label.

The categories under the weakly enhanced version of unlabelled employment psychological data are shown in formula (3).

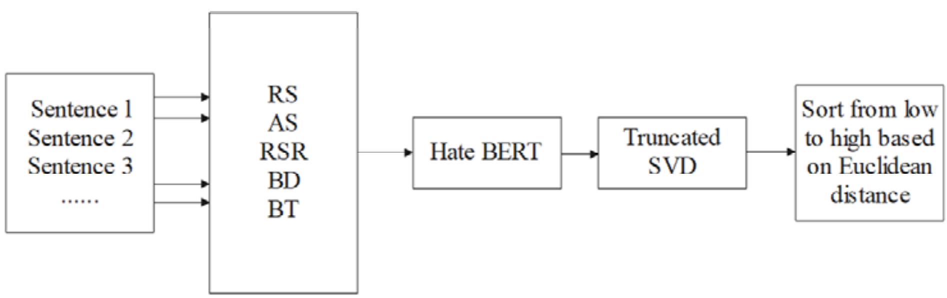
$$q_b = p_m\left(y \middle| \partial(u_b)\right) \tag{3}$$

Those greater than the set threshold are regarded as strong enhancement, and the cross-entropy loss is performed on the strongly enhanced samples of employment psychological data as shown in formula (4).

$$L_u = \frac{1}{\mu B}\sum_{b=1}^{\mu B} 1\left(\max\left(q_b \geq \tau\right)\right) H\left(q_b, p_m\left(y \middle| A(\mu_b)\right)\right) \tag{4}$$

Among them, τ is a scalar hyperparameter representing the threshold for retaining pseudo labels. The loss to be minimised is $L_S + \lambda_\mu L_\mu$, where λ_μ is a fixed scalar hyperparameter representing the relative weight of the unlabelled loss function.

Figure 2 Dynamic employment psychological data enhancement DAAS



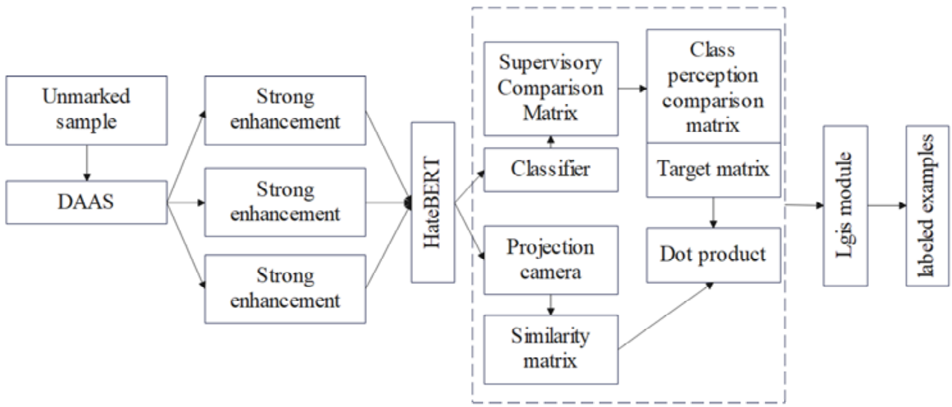
AS shown in Figure 2, the dynamic employment psychological data enhancement DAAS module performs five different employment psychological data enhancement methods for each sentence in the abusive language employment psychological dataset, including random exchange (RS), abstract summary (AS), random synonym replacement (RSR), random deletion (RD), and back translation (BT). Then, the enhanced samples of employment psychological data are encoded by HateBERT, and the Euclidean distance is calculated by truncating singular value decomposition. All samples are sorted from low to high based on distance, and the degree of enhancement of employment psychological data is distinguished by setting soft intervals.

In order to fully tap the potential information of employment psychological data and increase the diversity of training samples, the DAAS module strongly enhances the samples to the greatest extent without changing the semantics by truncating the singular value decomposition calculation, which can reduce the risk of overfitting of the model on the training set and improve the generalisation ability of the model. Its effectiveness will be verified in subsequent experiments.

3.2 Research on employment psychological data detection based on class perception and label-guided iterative optimisation

The composition of the Lgis-NewTextMatch framework is shown in Figure 3. The core modules are the green part of the Lgis module, the light yellow part of the semi-supervised module, and the pink part of the class-aware module.

Figure 3 Lgis-NewTextMatch framework diagram



Using the HateBERT model as the encoder, FC + Relu + FC + Sigmoid as the classifier, and the projection head as FC + Relu + FC, the employment psychological data enhancement module will enhance the DAAS with dynamic employment psychological data. In the class awareness module, only the strong enhancement methods filtered by DAAS are used.

Based on the TextMatch framework, inference prediction is performed to generate pseudo labels, and a supervised contrast matrix is constructed. Embeddings z from the same category is considered as positive pairs rather than different augmentations of the same sample.

However, it is unsafe to directly use the pseudo-label generated by TextMatch to construct the matrix, and the prediction of the model may contain a lot of noise. Therefore, contrastive learning for noise regularisation is introduced, T_{push} is used as the threshold for clustering and contrast in feature space, and then W_{scon} is converted into a class-aware contrast matrix W_{clacon}.

For the prediction of the model, if $p > T_{push}$, it is considered that the sample has a high probability of being in the distribution of employment psychological data U_{in}, and should be pulled closer with the same class. If $p < T_{push}$, the formula degenerates into contrastive learning, in which only strong enhancements in the same sample are positive samples. Each element W_{ij}^{clacon} in W_{clacon} has the following format (Thiem and Dasgupta, 2022):

$$W_{ij}^{clacon} = \begin{cases} 1, & \text{if } j = i \\ 1, & \text{if } z_i \text{ and } z_j \text{ come from the same category and} \\ & q_i \text{ and } q_j \text{ are both} > T_{push} \\ 0, & \text{other} \end{cases} \tag{5}$$

Among them, i and j are the indices of the augmented sample z. q_i represents the pseudo-label confidence score of the i^{th} augmented sample of employment psychology data.

Then, using the reweighting module, the class-aware contrast matrix is further processed, and the training of employment psychological data within the division with high confidence is further emphasised. At the same time, the foreground probability predicted by the model on weak enhancement is combined to generate the target matrix W_{target}, and each element W_{ij}^{target} of W_{target} is defined as shown in formula (6).

$$W_{ij}^{target} = \begin{cases} q_i \cdot q_j \cdot w_{ij}^{clacon} & \text{if } i \neq j \\ w_{ij}^{clacon} & \text{other} \end{cases} \tag{6}$$

The loss of the class-aware contrast module L_C is formulated by minimising the cross entropy between S (affinity matrix) and W_{target}.

The affinity matrix is obtained in the following ways: first, two strong enhancement methods are used to process the unlabelled employment psychological data, then the processed employment psychological data is input into the encoder $F(\cdot)$to extract features r, and then the projection head is used to map the features r to obtain a square-normalised low-dimensional embedding z.

Finally, the embedded affinity matrix $S \in \mathbb{R}^{2N \times 2N}$ is obtained by the dot product of the embedding $Z = \{z_i;\ i = 1, \ldots, 2N\}$ and the temperature factor τ, as shown in formula (7).

$$S_{ij} = \exp\left(z_i \cdot z_j / \tau\right) \tag{7}$$

The loss of the final loss L_C of the class-aware module Cacm is then defined as shown in formula (8).

$$L_C = H\left(S, W_{t\arg et}\right) = \sum_{i=1}^{2N} \frac{1}{1 + |P(i)|} L_{C,i} \tag{8}$$

Among them, H represents the cross entropy, $P(i)$ represents the prediction degree of other samples from the same category with confidence $p > T_{push}$. $|P(i)|$ represents its cardinality, and $|P(i)| + 1$ represents all positive pairs. $L_{c,i}$ is shown in formula (9).

$$-L_{c,i} = -\log \frac{\exp\left(z_i \cdot z_i^* / \tau\right)}{\sum_{i=1}^{2N} 1_{j \neq i} \exp\left(z_i \cdot z_j / \tau\right)} - \sum_{p \in P(i)} w_{ip} \cdot \log \frac{\exp\left(z_i \cdot z_p / \tau\right)}{\sum_{j=1}^{2N} 1_{j \neq i} \exp\left(z_i \cdot z_j / \tau\right)} \tag{9}$$

Among them, z_i^* is the embedding of the strong augmentation from the same sample instead of z_i. The final total loss is calculated using the weighted sum of the supervised contrast loss L_S, the semi-supervised loss L_u, and the class-aware contrast loss L_C, as shown in formula (10). L_X and L_u are the same loss functions used by TextMatch. λ_u and λ_C are the weights of the semi-supervised loss and the class-aware contrast loss, respectively.

$$L = L_S + \lambda_u L_u + \lambda_C L_C \tag{10}$$

The above is the class awareness part of the Lgis-NewTextMatch module, which uses pseudo-labels to form clusters that are beneficial to the classification task and at the same time uses contrastive learning to regulate and reduce the impact of 'out-of-distribution' employment psychological data. The class-aware module can combine the universality of self-supervised learning with the high efficiency of semi-supervised learning to achieve a more accurate and robust learning process.

Figure 4 Lgis strategy flowchart

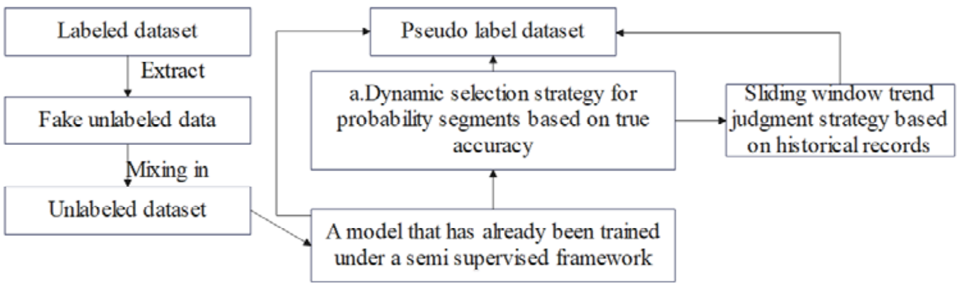
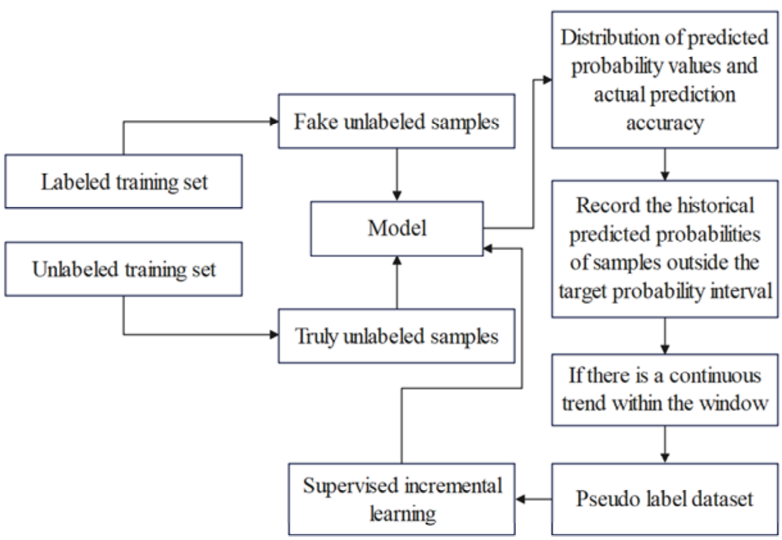


Figure 4 is the overall flow chart of Lgis strategy. Lgis is divided into two parts: probability segment dynamic selection strategy based on real accuracy and sliding window trend judgement strategy based on record history.

Figure 5 Flowchart of sliding window trend judgement based on history record



For the interval segments that were not selected in the previous process, a queue Q will be created and maintained for them. This queue records the model output probability $Q = \{q_1, q_2, \ldots, q_n\}$ each time, and n represents the number of times the sample is currently recorded.

Then, based on this queue, a sliding window of length $size_W$ is constructed. As shown in Figure 5, it is determined whether the change trend of the probability output of the model for the sample within $size_W$ times $size_W$ (when the length of Q is greater than or equal to $size_{W+1}$) increases (namely, whether $q_n - q_{n-1} > 0$) or decreases. If the $size_W$ times of change trends in the window are completely consistent, or the proportion of a certain trend is greater than a threshold t_W, then the correct label of the model is considered to be the label that follows the direction of this trend. In other words, if the model output probability of a sample always rises in the $size_W$ time records and is greater than 0.5 (the middle point of the y value of sigmoid) or a certain set threshold t_{p1}, then its pseudo label is considered to be 1. On the contrary, if it always decreases in the records and is less than t_{p2}, then its pseudo label is considered to be 0.

In this experiment, $t_{P1} = 0.6$ and $t_{P2} = 0.4$ are set, and a margin width of 0.2 is obtained. At the same time, $size_w = 4$ and $t_w = 0.74$ are set, that is, when 3/4 of the history shows that the current sample has the same trend, it will actively promote this trend. High confidence is defined as the model output probability value p is closer to 0 or 1, that is, abs(1 – p) is smaller, while low confidence is closer to the middle value 0.5, that is, abs(0.5 – p) is smaller.

4 Test study

4.1 Test methods

The validated multidimensional mental health literacy questionnaire is adopted, which includes five dimensions: symptom identification, professional willingness to seek help, treatment compliance, resource information acquisition, and attitude towards patients with mental illness. The symptom recognition dimension is evaluated in the form of true or false questions. Each question has three options: correct, wrong, and do not know. If the subject answers the question correctly, 1 point will be awarded, and if the subject answers incorrectly or selected ‘I do not know’, 0 point will be awarded. The remaining dimensions are rated on a 6-rating scale, from 1 (strongly disagree) to 6 (strongly agree). The higher the score of the questionnaire, the better the mental health literacy. In this survey, the Cronbach’s α coefficient of Likert scoring question is 0.771, and the Bank-Richardson coefficient of judgement question is 0.659.

This study uses the proposed model to statistically analyse the influencing factors of college students’ employment psychology. The specific analysis ideas are: first, this paper conducts descriptive statistics, correlation analysis and regression analysis on the relevant variables and variable dimensions involved in the study. Then, this paper uses latent profile analysis to divide and name the types of graduates’ career anxiety. Finally, this paper uses the bootstrap method to test the chain mediation effect of psychological resilience and coping style.

The experimental hardware environment configuration is shown in Table 1.

Using the dataset HS&OL as the experimental dataset, this paper explores the advantages of semi-supervised method in low-resource situations, and randomly selects 10,000 labelled data from the dataset to use fully supervised learning. Then, we randomly select 20%, 40%, 60% and 80% of labelled data for each dataset for semi-supervised learning.

The model of this paper is Lgis-NewTextMatch, and the comparison algorithms are random guess, Π-model, MeanTeacher, PseudoLabel, VAT, MixMatch, UDA, ReMixMatch, FixMatch. The performance of this model is compared and analysed.

Table 1 Experimental environment and configuration

OS	Ubuntu 22.04 LTS
CUDA	11.8
GPU	NVIDIA GeForce RTX 4090 24 G
pytorch	2.1.2
transformers	4.24
lightning	2.1.2
Python	3.8.17

4.2 Results analysis and discussion

4.2.1 Model performance test

The model is reproduced in this experimental environment. The experimental results are shown in Table 2.

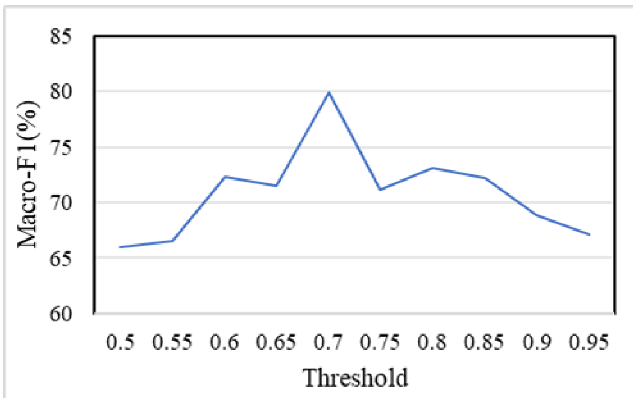
Table 2 Comparative experimental results

Full supervision		77.85		
Marking data ratio	20%	Marking data ratio	20%	Marking data ratio
Random guess	45.02	Random guess	45.02	Random guess
Π-Model	52.99	Π-Model	52.99	Π-Model
MeanTeacher	61.70	MeanTeacher	61.70	MeanTeacher
PseudoLabel	55.05	PseudoLabel	55.05	PseudoLabel
VAT	58.24	VAT	58.24	VAT
MixMatch	64.17	MixMatch	64.17	MixMatch
UDA	63.38	UDA	63.38	UDA
ReMixMatch	67.26	ReMixMatch	67.26	ReMixMatch
FixMatch	68.32	FixMatch	68.32	FixMatch
Lgis-NewTextMatch	71.93	Lgis-NewTextMatch	71.93	Lgis-NewTextMatch

From the comparison of experimental results in Table 2, it can be seen that the performance of the Lgis-NewTextMatch framework is better than that of the baseline model. Lgis-NewTextMatch selects more reasonable weak enhancement and strong enhancement for unlabelled data through DAAS module, which avoids the noise introduced into the model due to fixed or combined data enhancement in the field of employment psychological factors. At the same time, compared with the original data, reasonable weak enhancement provides a certain data disturbance and helps the model better mine the potential characteristics of the data. Secondly, the DAAS module can improve the generalisation of the model by using the maximum strong enhancement without changing the semantics. Lgis-NewTextMatch implements supervised comparative learning on labelled data, which makes the model better consider inter-class information in low-resource situations. Therefore, Lgis-NewTextMatch shows better performance. Compared with FixMatch, Lgis-NewTextMatch has 20% labelled data. Moreover, it increased by 1.53%, 1.58% on 40% labelled data, 1.62% on 60% labelled data, and 1.57% on 80% labelled data. Lgis-NewTextMatch only uses 60% of the labelled data, which is only 1.91% worse than the full supervision effect. In addition, after using 80% of the labelled data, it is only 0.16% worse than the full supervision effect.

Since TextMatch determines whether the label generated by weak enhancement can be used as a pseudo label by setting a threshold τ, 20% of the label data is randomly selected for the HS&OL dataset to explore the influence of different threshold τ settings on the experimental results. The results are shown in Figure 6.

Figure 6 Influence of threshold τ on results (see online version for colours)



As shown in Figure 6, as the threshold value τ increases, the performance of the model increases initially and reaches its highest point at 0.65. At this point, the model achieves the best matching performance. Due to the increase of threshold τ, the model filters some false labels and fully learns high-confidence data features, which improves the accuracy and stability of the model. As the threshold continues to increase, the model fluctuates slightly. The reason is that some samples with large noise or blurred boundaries are encountered, which cause the performance to decline first. Then, with the addition of more high-quality samples, the model performance recovers until the threshold reaches 0.8. After the threshold reaches 0.8, it has been showing a downward trend. The reason is that the threshold is too large, which makes the data samples that match the pseudo-labels of weak enhancement predictions fewer and fewer. Although to a certain extent, the accuracy of the pseudo-labels may be more accurate and reduce a certain degree of noise, too little sample data affects the generalisation ability of the model, resulting in poor model performance. To sum up, the setting of the threshold has a significant impact on the performance of Lgis-NewTextMatch. There is an optimal range that needs to be explored through experiments, and the model can obtain the best performance within this range. However, after exceeding this range, the performance will start to decline because the threshold is too large, indicating that a balance point needs to be found when adjusting the threshold to ensure that the model has sufficient flexibility while maintaining good generalisation ability.

4.2.2 Analysis of the influencing factors of college students' employment psychology

To achieve the standard of mental health literacy, the following conditions need to be met at the same time: the total score of symptom recognition dimension is greater than or equal to 7 points, and the total average score of resource information acquisition, professional willingness to seek help, and treatment compliance dimension is all greater than or equal to 4 points. In this paper, the compliance rate and average score of each dimension are counted in Table 3.

Table 3 Compliance rate of each dimension of mental health literacy

Dimension	*Score*	*Compliance rate*
Treatment compliance	4.68	86.81%
Resource information acquisition	4.64	85.55%
Professional willingness to help	4.49	84.02%
Attitudes towards people with mental illness	4.46	76.15%
Symptom recognition	6.68	67.78%

Table 4 Descriptive statistics of each dimension of career choice anxiety

Dimension	*M*	*SD*
Employment competitive pressure	2.68	98.01%
Lack of employment support	2.61	92.07%
Lack of self-confidence	2.63	97.02%
Employment outlook worries	2.13	71.28%
Career choice anxiety	2.60	86.13%

The descriptive statistics of each dimension of career choice anxiety are shown in Table 4. The average score of the respondents in this study on career choice anxiety is 2.63, and the standard deviation is 0.87. The average and standard deviation of each dimension are shown in the following table. Although the overall score of the graduates surveyed in this research did not reach the theoretical value of 3 points, 36.41% of the graduates still had obvious anxiety about choosing a job. Because there are four sub-dimensions of career choice anxiety, the calculation method of average score ignores the performance of different groups in different dimensions of career choice anxiety, which is not the best way to describe graduates' career choice anxiety. Therefore, in order to explore the types of graduates' career choice anxiety and the influence of demographic variables, this paper uses the model to analyse the potential profile of graduates' career choice anxiety.

First, a baseline model is established, and then the number of potential categories is gradually increased, and the models are compared one by one until the best fitting model is found. The indices for evaluating the quality of model fitting mainly include the information index Akaike information criterion (AIC), Bayesian information criterion (BIC), adjusted Bayesian information criterion (aBIC) and entropy based on the log-likelihood function value (loglikelihood, LL). The smaller the values of AIC, BIC and aBIC, the better the model fits. The higher the value of entropy, the more accurate the model classification is. When the entropy is above 0.8, more than 90% of the data are accurately classified, and the p-values of LMR and BLRT are significant, indicating that k-class model fits are better than those of k – 1 class models.

The fitting information of potential profile analysis of career choice anxiety is shown in Table 5. The average attribution probabilities (columns) of each potential category of research subjects (rows) are shown in Table 6.

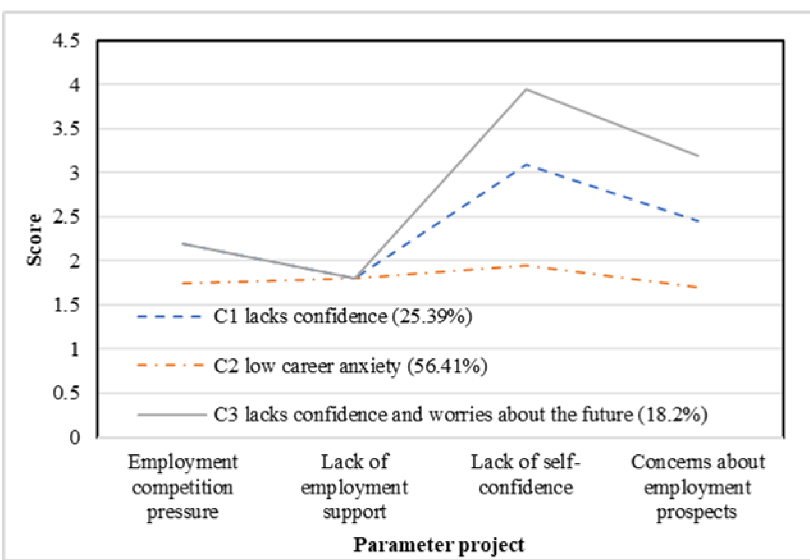
Table 5 Fitting information of potential profile analysis of career choice anxiety

Category	*Likelihood ratio*	*Entropy*	*AIC*	*BIC*	*aBIC*	*LMR (p)*	*BLRT (p)*	*Proportion of each category (%)*
1	–1,646.05	-	3,307.95	3,339.36	3,314.22	-	-	-
2	–1,324.12	0.93	2,673.97	2,725.01	2,684.18	$p < 0.001$	$p < 0.001$	59.19/40.81
3	–1,221.15	0.92	2,477.93	2,548.61	2,492.07	$p < 0.001$	$p < 0.001$	25.41/56.39/ 18.19
4	–1,185.54	0.90	2,416.63	2,506.94	2,434.69	0.11	$p < 0.001$	53.79/17.71/ 21.31/7.21

Table 6 Average belonging probabilities (columns) of study subjects (rows) in each potential category

	C1 (%)	*C2(%)*	*C3 (%)*
C1	93.26	2.14	4.6
C2	2.29	97.71	0
C3	5.7	0	94.3

Figure 7 Analysis of potential profiles of three categories of career choice anxiety (see online version for colours)



In Tables 5 and 6, with the increase of the number of categories, the values of AIC, BIC and aBIC gradually decrease and are always above 0.9. However, when the model has four potential categories, the p = 0.11 > 0.05 of LMR, indicating that there is no significant difference between the 4-category model and the 3-category model, and the minimum category of the 4-category model only accounts for 7.52%, so the simpler 3-category model is finally selected. The attribution probability of the three categories is between 93.26-97.71%, indicating that the potential classification of the three categories is credible.

The scores of the three potential categories in the four dimensions of employment competition pressure, lack of employment support, insufficient self-confidence and concern about employment prospects are shown in Figure 7. Individuals in category 1 score lower in the dimensions of employment competition pressure, lack of employment support and worry about employment prospects, but score higher in the dimension of lack of self-confidence, so they are named as lack of self-confidence. Individuals in category 2 score lower in four dimensions of career choice anxiety, so they are named as low career choice anxiety type. Individuals in category 3 score lower in the two dimensions of employment competitive pressure and lack of employment support, but score higher in the dimensions of lack of self-confidence and worry about employment prospects, so they are named as lack of self-confidence and worry about prospects.

Table 7 Scores of three categories of career anxiety in four dimensions

Dimension of career anxiety	*C1 Lack of self-confidence*	*C2 Low career choice anxiety*	*C3 Lacks confidence and worries about the outlook*	*F*	$\eta^2 p$
	M ± SD	*M ± SD*	*M ± SD*		
Employment competitive pressure	2.17 + 0.05	1.73 + 0.03	2.16 + 0.05	899.94*** 3 > 1 > 2	0.819
Lack of employment support	1.79 + 0.07	1.84 + 0.030	1.69 + 0.08	798.11*** 3 > 1 > 2	0.899
Lack of self-confidence	3.15 + 0.07	1.93 + 0.03	3.99 + 0.06	650.97*** 3 > 1 > 2	0.769
Employment outlook worries	2.49 + 0.05	1.59 + 0.03	3.26 +0.06	649.79*** 3 > 1 > 2	0.769

Note: * indicates p < 0.05, * * indicates p < 0.01, * * * indicates p < 0.001

Taking the above three types of career choice anxiety as grouping variables, one-way ANOVA is carried out on the four dimensions of career choice anxiety, and the results are shown in Table 7. The three types of career choice anxiety have significant differences in each dimension, and the results of post-hoc analysis show that there are significant differences in the scores of the three types of potential career choice anxiety in the four dimensions.

The mental health literacy, psychological resilience, positive and negative coping styles, and career choice anxiety of fresh graduates are analysed. The Pearson product difference correlation coefficient is shown in Table 8. There is a significant correlation between mental health literacy, positive and negative coping styles, and career choice anxiety.

Table 8 Results of correlation analysis

		1	*2*	*3*	*4*	*5*
1	Mental health literacy	1				
2	Psychological resilience	0.26	1			
3	Actively respond	0.24	0.62	1		
4	Negative coping	–0.22	–0.38	–0.31	1	
5	Career choice anxiety	–0.29	–0.60	–0.52	0.55	1

Note: * indicates $p < 0.05$, * * indicates $p < 0.01$, * * * indicates $p < 0.001$.

When graduates with higher mental health literacy face stressful events such as difficulties and setbacks in career choice, their mental health knowledge, service resources and active and cooperative attitude of seeking help enable them to quickly adapt to difficult situations and recover quickly. Graduates with higher psychological flexibility can better mobilise their internal and external resources. Therefore, when faced with difficulties in choosing a career, they can alleviate their anxiety in choosing a career by adopting positive coping methods such as solving problems, collecting information about choosing a job, making preparations for choosing a job and seeking help appropriately. On the contrary, when graduates with low mental health literacy level are faced with difficulties and setbacks in career choice, their poor mental health knowledge and resources make them deeper and deeper in difficult situations, and it is difficult to integrate their own resources to actively and effectively cope with employment pressure. Instead, they use more negative coping styles such as self-blame, evasion and neglect, so they feel more anxious.

The negative coping style model can always explain more career choice anxiety than the positive coping style model. On the one hand, the low reliability of internal consistency of positive coping styles may affect the explanation ratio of the total variation of career choice anxiety in regression analysis.

Improving graduates' mental health literacy can directly improve their mental health knowledge and attention to mental health, and indirectly improve individual mental health status, but it is difficult to directly increase graduates' social and environmental resources during career selection, so the buffering effect of positive coping styles will be limited by individual coping resources. However, negative coping style is an individual's self-defence behaviour in the face of stressful events. Compared with positive coping style, it is less likely to be restricted by individual environmental resources. Therefore, graduates with low mental health literacy have poor mental health knowledge, and they can take more avoidance, denial, fantasy, etc. to cope with the pressure in career choice according to their own wishes. In addition, excessive use of negative coping style may aggravate anxiety.

5 Conclusions

In order to further improve the analysis effect of data on factors affecting college students' employment psychology, this paper introduces the class-aware comparative learning Cacm module and the label-guided iterative self-incremental learning Lgis module to help the model fully tap the potential characteristics of the data in unlabelled data, and at the same time effectively solve the problem of insufficient labelled data on factors affecting college students' employment psychology. The class awareness module can improve the quality of pseudo labels, reduce the noise contained by artificial labels generated by itself, and emphasise the learning of 'clean labels'. Moreover, this paper verifies the effectiveness of the Lgis-NewTextMatch model by combining comparative experiments, and applies the model to the analysis of psychological factors affecting college students' employment. Through analysis, it can be seen that graduates with higher mental health literacy have higher levels of psychological resilience. When encountering difficulties in the career selection process, they tend to use more positive and mature coping methods and less negative and immature coping methods, which can buffer the career selection anxiety caused by career selection difficulties. When providing employment guidance, schools should work in stages and groups. Therefore, when providing career guidance, schools should carry out their work in stages and groups, cultivate students' psychological resilience and positive coping methods in the face of setbacks, and enhance graduates' self-confidence in their own career development.

Although the data enhancement methods in this paper are effective in selecting different data enhancement methods, they are all based on the processing of original data rather than the transformation of vector dimensions. Therefore, the dynamic selection method based on vector dimension can be studied in the future to verify whether it is still effective.

All the variables in this study come from the self-report method. Although the common method has little deviation, the self-report method cannot accurately measure the participants' implicit attitude towards patients with mental illness, and it can be combined with implicit association test and other methods in the future. Secondly, when verifying the multi-dimensional mental health literacy questionnaire, due to the low reliability of understanding risk factors, this dimension is directly discarded. In the future, this dimension can be supplemented with questions or integrated with the symptom recognition dimension. For example, it can measure whether individuals can find hidden dangers of mental health problems in daily life and avoid them correctly.

Declarations

The author declares that he has no conflicts of interest.

References

Alshurideh, M.T., Al Masaeid, T., Alzoubi, H.M., El Khatib, M., Ghazal, T.M., Hamadneh, S., … and Sattar, O. (2022) 'Components determining the behavior and psychological impact of entrepreneurship among higher vocational students', *Journal for ReAttach Therapy and Developmental Diversities*, Vol. 5, No. 2s, pp.189–200.

Apriceno, M., Lytle, A., Monahan, C., Macdonald, J. and Levy, S.R. (2021) 'Prioritizing health care and employment resources during COVID-19: roles of benevolent and hostile ageism', *The Gerontologist*, Vol. 61, No. 1, pp.98–102.

Baluku, M.M., Mugabi, E.N., Nansamba, J., Matagi, L., Onderi, P. and Otto, K. (2021) 'Psychological capital and career outcomes among final year university students: the mediating role of career engagement and perceived employability', *International Journal of Applied Positive Psychology*, Vol. 6, No. 1, pp.55–80.

Belle, M.A., Antwi, C.O., Ntim, S.Y., Affum-Osei, E. and Ren, J. (2022) 'Am I gonna get a job? Graduating students' psychological capital, coping styles, and employment anxiety', *Journal of Career Development*, Vol. 49, No. 5, pp.1122–1136.

Charoensukmongkol, P. and Phungsoonthorn, T. (2022) 'The interaction effect of crisis communication and social support on the emotional exhaustion of university employees during the COVID-19 crisis', *International Journal of Business Communication*, Vol. 59, No. 2, pp.269–286.

Easterbrook, M.J. and Hadden, I.R. (2021) 'Tackling educational inequalities with social psychology: identities, contexts, and interventions', *Social Issues and Policy Review*, Vol. 15, No. 1, pp.180–236.

Kee, C.E. (2021) 'The impact of COVID-19: graduate students' emotional and psychological experiences', *Journal of Human Behavior in the Social Environment*, Vol. 31, Nos. 1–4, pp.476–488.

Kotera, Y., Green, P. and Sheffield, D. (2022) 'Positive psychology for mental wellbeing of UK therapeutic students: relationships with engagement, motivation, resilience and self-compassion', *International Journal of Mental Health and Addiction*, Vol. 20, No. 3, pp.1611–1626.

LaCosse, J., Murphy, M.C., Garcia, J.A. and Zirkel, S. (2021) 'The role of STEM professors' mindset beliefs on students' anticipated psychological experiences and course interest', *Journal of Educational Psychology*, Vol. 113, No. 5, pp.949–961.

Liu, X.Q., Guo, Y.X. and Xu, Y. (2023) 'Risk factors and digital interventions for anxiety disorders in college students: stakeholder perspectives', *World Journal of Clinical Cases*, Vol. 11, No. 7, pp.1442–1453.

Ma, Y., Huang, G. and Autin, K.L. (2021) 'Linking decent work with academic engagement and satisfaction among first-generation college students: a psychology of working perspective', *Journal of Career Assessment*, Vol. 29, No. 1, pp.148–163.

Mascherini, G., Catelan, D., Pellegrini-Giampietro, D.E., Petri, C., Scaletti, C. and Gulisano, M. (2021) 'Changes in physical activity levels, eating habits and psychological well-being during the Italian COVID-19 pandemic lockdown: impact of socio-demographic factors on the Florentine academic population', *PLOS ONE*, Vol. 16, No. 5, pp.e0252395–e0252406.

Mbous, Y.P.V., Mohamed, R. and Rudisill, T.M. (2024) 'International students' challenges during the COVID-19 pandemic in a university in the United States: a focus group study', *Current Psychology*, Vol. 43, No. 9, pp.8325–8337.

Moriarty, T., Bourbeau, K., Fontana, F., McNamara, S. and Pereira da Silva, M. (2021) 'The relationship between psychological stress and healthy lifestyle behaviors during COVID-19 among students in a US Midwest university', *International Journal of Environmental Research and Public Health*, Vol. 18, No. 9, pp.4752–4761.

Peltz, J.S., Bodenlos, J.S., Kingery, J.N. and Rogge, R.D. (2021) 'The role of financial strain in college students' work hours, sleep, and mental health', *Journal of American College Health*, Vol. 69, No. 6, pp.577–584.

Phungsoonthorn, T. and Charoensukmongkol, P. (2022) 'How does mindfulness help university employees cope with emotional exhaustion during the COVID-19 crisis? The mediating role of psychological hardiness and the moderating effect of workload', *Scandinavian Journal of Psychology*, Vol. 63, No. 5, pp.449–461.

Prasath, P.R., Mather, P.C., Bhat, C.S. and James, J.K. (2021) 'University student well-being during COVID-19: the role of psychological capital and coping strategies', *Professional Counselor*, Vol. 11, No. 1, pp.46–60.

Quintiliani, L., Sisto, A., Vicinanza, F., Curcio, G. and Tambone, V. (2022) 'Resilience and psychological impact on Italian university students during the COVID-19 pandemic', *Psychology, Health & Medicine*, Vol. 27, No. 1, pp.69–80.

Ryu, S. and Fan, L. (2023) 'The relationship between financial worries and psychological distress among US adults', *Journal of Family and Economic Issues*, Vol. 44, No. 1, pp.16–33.

Thiem, K.C. and Dasgupta, N. (2022) 'From precollege to career: barriers facing historically marginalized students and evidence-based solutions', *Social Issues and Policy Review*, Vol. 16, No. 1, pp.212–251.

Zapata-Cuervo, N., Montes-Guerra, M.I., Shin, H.H., Jeong, M. and Cho, M.H. (2023) 'Students' psychological perceptions toward online learning engagement and outcomes during the COVID-19 pandemic: a comparative analysis of students in three different countries', *Journal of Hospitality & Tourism Education*, Vol. 35, No. 2, pp.108–122.