Assessment of the Mental Health of Students Using Neural Networks

Chu Yanhua
Henan Vocational College of Art
Zhengzhou, Henan, China
ewvrn2887619@163.com

ABSTRACT: *With the increase of social pressure and mental health problems, evaluating and diagnosing college students' mental health has become particularly important. To improve the accuracy and efficiency of assessment and diagnosis, we propose a neural network-based mental health model for college students. This model adopts deep learning technology to assess and diagnose students' mental health by analyzing their personal information, behavioral data, and academic performance. Specifically, we used models such as Convolutional Neural Networks (CNN) and Long Short-Term Memory Networks (LSTM) to conduct a multidimensional analysis of students' learning behavior, emotional changes, and social situations. At the same time, we have also utilized advanced technologies such as self-attention and attention mechanisms to better capture students' information from different dimensions. Through extensive experimental data validation, we found that the neural network-based mental health model for college students has high accuracy and efficiency. This model has higher diagnostic accuracy and a lower misdiagnosis rate than traditional evaluation and diagnostic methods. In addition, the model also has good generalization performance and can adapt to the application needs of different types of students and different scenarios.*

Keywords: College students' Mental Health, Neural Network Prediction Model, College students' Psychological Research, Neural Network

Received: 3 March 2023, Revised 19 June 2023, Accepted 28 June 2023

DOI: 10.6025/jnt/2023/14/3/68-75

1. Introduction

The mental health and well-being of undergraduates are crucial to their growth and an important contribution to the community. There is evidence that mental health problems are fragile in a university context [1]. Many studies have shown that undergraduates worldwide have a high level of stress and depression. With the gradual increase in college students, they have also maintained good health. Still, in the face of intense competition among college students, study pressure, improper interpersonal relationships, family environment, future employment and other factors, the psychological problems of adolescents gradually increased; more than 20% of young people experience mental disorders at some point in their lives [2]. For young people, the transition to college can be daunting, which means they can get unprecedented freedom from hometown and away from a long list of homework and exams. Because of these factors, college students are considered particularly vulnerable to depression [3]. A study in 2010 found that, among the young people during the Great Depression, more than five times of high school and college students have anxiety and other mental health problems. Their psychological burden is much larger than we think, and their psychological problems are inevitable especially depression and anxiety. Depression and anxiety are two of college students' most common psychological problems [4]. Depression is a kind of depressed emotional state, leading to diminished interest or even loss; people often blame themselves and mental fatigue, and there is a strong sense of helplessness; they become solitary and are not good at dealing with people when faced with difficulties, they are alone and do not feel the help of the community, they cannot deal with the difficulties brought about by all aspects, become confused, and deny themselves [5]. In recent years, the proportion of depression in college students has continued to rise, and severe depression has a particularly serious impact on the university, such as learning difficulties, barriers to communication, evasion of difficulties, and even suicidal behaviour and thoughts. Under normal circumstances, if a period, the mood of depression is not stopped, it will be transformed into depression. Depression is characterized by irritable mood, sleep disorder, and somatic complaints. According to the survey, more than 29% of South Korea and United States college students have reported depression problems, and depression can lead to social problems or suicide. Indeed, most suicides are from adolescents [6].

2. Material and Methods

2.1. Extreme Learning Machine Algorithm

Extreme learning machine algorithm is a single layer feedforward neural network algorithm. The ELM model diagram of the single hidden layer is shown in Figure 1.

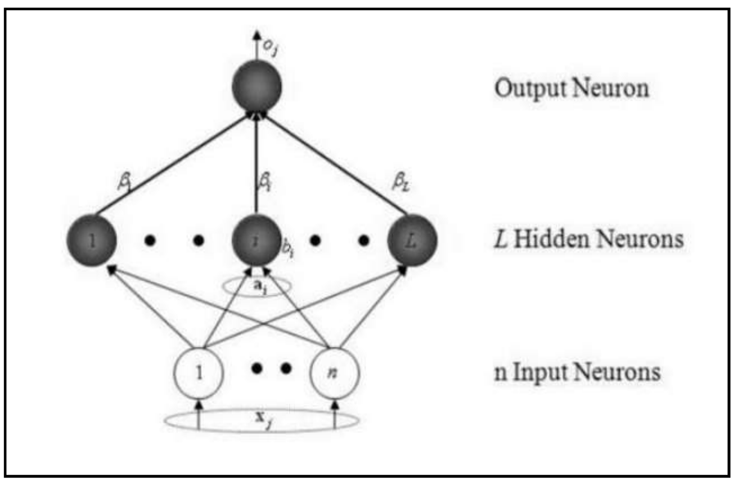

Figure 1. ELM model of single hidden layer

For a single hidden layer neural network, assume that N arbitrary samples (X_i, Y_i) have:

$$X_i = [x_{i1}, x_{i2}, ..., x_{in}]^T \in R^n \quad (1)$$

$$Y_i = [y_{i1}, y_{i2}, ..., y_{in}]^T \in R^n \quad (2)$$

Sequences X_i and Y_i denote the input sequence and output sequence. A single hidden layer neural network with L hidden nodes is represented as follows:

$$\sum_{i=1}^{L} \beta_i g(W_i g X_j + b_j) = o_j \quad (3)$$

Among them, $g(g)$ is the activation function, $W_i = [w_{i1}, w_{i2}, ..., w_{im}]$ is the input weight, the offset of the i^{th} hidden layer is denoted by b_i, and $W_i g X_j$ the inner product of the two. The ultimate goal of single hidden layer neural network learning is to minimize the output error, which is expressed as:

$$\sum_{j=1}^{L} \|o_j - y_j\| = 0 \quad (4)$$

So, there is β_i, W_i, b_i, so that:

$$\sum_{j=1}^{L} \beta_i g(W_i g X_j + b_j) = y_j \quad (5)$$

The matrix form of the above equation can be expressed as: $H\beta = Y$. Among them: H represents the output of the hidden layer node, Y is the expected output. In order to be able to train a single hidden-layer neural network so that the error is minimized, this paper expects the resulting $\hat{W}_i$, b_i, $\hat{\beta}_i$ to:

$$\|H(\hat{W}_i, b_i)\hat{\beta}_i - Y\| = \min_{W,b,\beta} \|H(W_i g b_i)\beta_i - Y\| \quad (6)$$

Here, $i = 1,2,......K$ is equivalent to minimizing the loss function

$$E = \sum_{j=1}^{N} (\sum_{j=1}^{L} \beta_i g(W_i g X_j + b_i) - y_i)^2 \quad (7)$$

In the ELM algorithm, once the input weights and offsets are determined, the output matrix H of the hidden layer will be uniquely determined, so the training of a single hidden layer neural network can be transformed into solving a linear system $H\beta = Y$. Then the output weight can be determined as:

$$\hat{\beta} = H^{\dagger} T \quad (8)$$

Here, H^+ is the Moore-Penrose generalized inverse of matrix H, and it can also be proved that the singular number of $\hat{\beta}$ is the smallest and unique. ELM algorithm has the advantages of fast learning speed and generalization ability, but the ELM algorithm has poor stability. In predictive engineering, the ELM prediction can be used to predict the average of the predicted results as a predictive result. This is called the Extended Extreme Learning Machine Algorithm. EELM is also used to predict the experimental data in this paper [7].

2.2. Neural Network

The neural network is called artificial neural network, which originates from the study of biological neurons and is an intelligent computer network system simulating biological neural networks. Like the human brain, neural networks consist of interconnected neurons called processing units or nodes. Nodes are used to process, remember certain information, and work in parallel with other nodes. The structure and function of each node is relatively simple, but the system composed by a large number of nodes is very complex. The connection between nodes is called edge, which can reflect the strength of relevance through the size of weight. The weights of the connecting nodes can be changed to train the neural network to achieve the function.

Structural model of neurons. As the most basic processing unit of neural network, neurons are generally multi-input and single-output nonlinear components. Its general structure model is shown in Figure 2.

Here, U_i is the state inside neuron i, θ_i is the threshold, and X_j is the input signal. W_{ij} represents the weight connected to neuron X_i, and S_i represents an externally input control signal. A complete neuron or node must include adders and activation functions, and the adder and activation functions are described in detail below:

Adder: The input of node is represented by X, the output of node is denoted by Y, the weight of neural network connected with the upper layer is represented by W, represents the deviation of node. Then the adder U_j of the *jth* node can be defined as:

$$U_j = \sum_{j=1}^{n} W_{ij} X_i + \theta_i \qquad (9)$$

Among them, n is the number of input nodes in the upper layer and X_i is the output of the i^{th} node in the upper layer. Since the output of each upper node is used as the input of this node, there are n inputs in total. W^{ij} represents the weight of the i^{th} node in the upper layer connected to the j^{th} node in this layer. θ_i can be regarded as a constant term or a threshold value. The purpose of the adder is to do a linear combination of the inputs in this layer. During the operation of the adder, the most crucial determination of the weight coefficient W^{ij} is made.

Activation function: The activation function of the j^{th} node is defined as:

$$Y_j = f(U_j) \qquad (10)$$

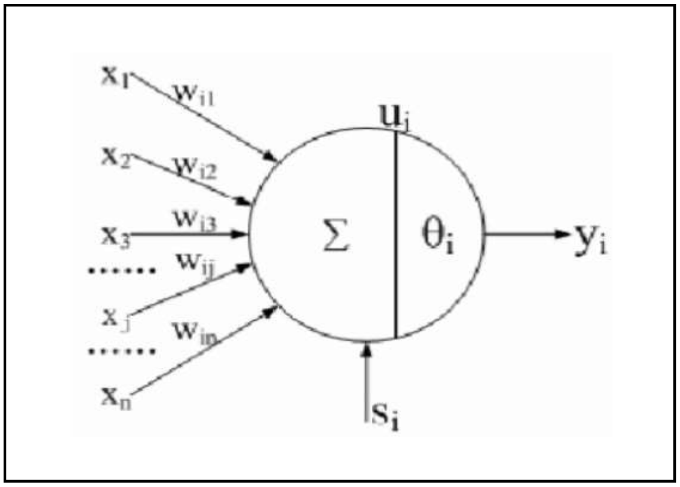

Figure 2. Neuron model

Among them, Y_j is the activation function value, which is also the output of the node. $f(g)$ indicates that the parameter is an activation function of the adder jU. Generally, has the following types: a) [0,1] jump function; b) [-1 , 1] jump function; c) (0,1) Sigmoid function; d) (-1,1) Sigmoid function. The specific graphical representation of the Sigmoid function is shown in Figure 3.

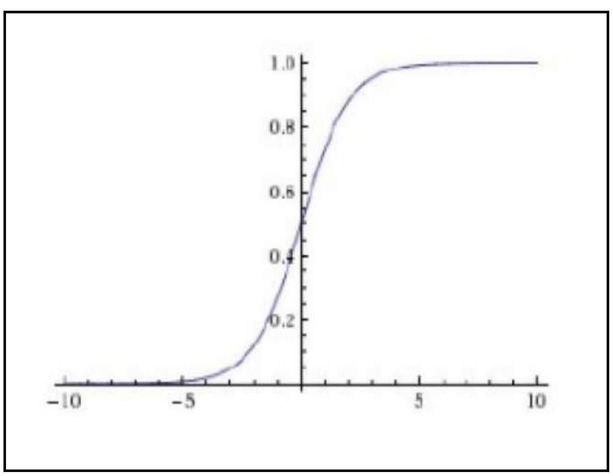

Figure 3. (0, 1) Sigmoid function

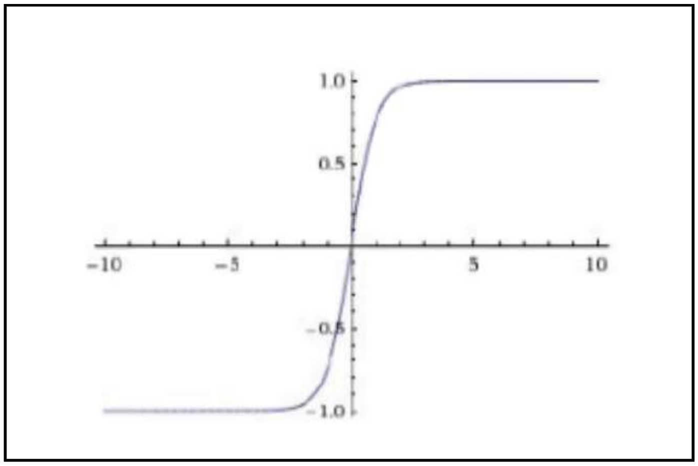

Figure 4. (-1,1) Sigmoid function

Activation functions usually satisfy the following properties: a) Nonlinear. b) Differentiability: This property must be satisfied when the optimization method is based on gradient. C) Monotonicity: When the activation function is a monotonous function, the single-layer network can be guaranteed to be convex function. d) When the activation function satisfied $f(x) \approx x$, if the parameter initialization value is small, the neural network has a higher training efficiency; conversely, if the activation function does not meet $f(x) \approx x$, then it needs to be careful to set the initialization value. In the case of an infinite range of activation functions, model training is more efficient, but in this case, it is generally necessary to set a smaller learning rate. Under the action of adders and activation functions, nodes play the role of equivalent to a hyperplane. In the ideal case, the points above the hyperplane belong to one category, the points on the other side of the hyperplane belong to another category, and the points in different regions belong to different categories so that the prediction of the multi-category problem can be finally realized.

2.3. General Modeling Steps of Neural Networks

The establishment of neural network generally includes: data preparation, network structure determination, network weight determination three steps:

• **Data Preparation Phase:** Neural network input variables are usually required in the 0-1 range to guarantee the input variables in the same dimension and improve the accuracy of prediction, it is important to train the neural network.

• **Determination of Network Structure:** In general, the number of neural network layer and the number of nodes in each layer determine the complexity of the neural network structure, among them, the number of hidden layers and the number of hidden nodes are particularly critical, the more the number of the two, the more complex the neural network. The convergence rate for neural networks with few hidden layers is fast in learning, but the prediction accuracy is low. Neural networks with many layers have high predictive accuracy but may have the possibility of non-convergence. Therefore, the network training complexity and model training efficiency need to be weighed in determining the number of layers. Although theoretically, multi-hidden neural networks imply higher accuracy, in reality, multi-hidden neural networks may not necessarily get the optimal solution and are more complicated. Therefore, in the actual prediction, it is often the most effective to use a single hidden layer neural network to predict. There is no clear standard for the number of hidden nodes, but too many in the modelling process often lead to over-fitting.

• **Determination of Neural Network Weights:** A neural network explores the complex relationship between input variables and output variables through proper network structure and finds the quantitative relationship and regularity between the input variables and the output variables through repeated analysis and study of the existing variables and embodies the law in the weights. The determination of the weights generally includes the following steps:

(1) Weight Initialization. In neural networks, weights are typically initialized to a random number between -0.5 and 0.5. The reason for initializing weights to approximately 0 is that the Sigmoid activation function approximates the neural network to a linear model, and then as the weights increase, the turns model again to a non-linear model;

(2) calculate the adder and activation function values of each unit to obtain the predicted value of the sample;

(3) compare the error between the predicted value and the actual value of the sample, and adjust the weight of each network according to the error;

(4) return to (2); the adjustment and calculation of weights are iterative processes until the error between the final predicted value and the actual value reaches a relatively small range. When the weights are set, the hyperplane is also determined.

3. Results

3.1. Diagnostic Model of Mental Health

In order to determine the relationship between the tendency of psychological problems and the causes of the problems, this paper evaluates the mental health status of college students through the questionnaire test: 90 symptom checklist (SCL-90), and makes 1-5 grades, any factor score or positive average are greater than or equal to 3, and there are mental health problems. If anyone in 9 factors ≥ 3 points or positive average ≥ 3 points, the detection rate of psychological problems was 12.77%, mainly distributed in 3.61% of hostile emotions, 3.38% of obsessive-compulsive symptom, 3.3% of interpersonal sensitivity, and 2.68% of paranoid symptoms. This article selects four typical mental health problems in college students for analysis according to their different reasons.

The set of psychological problems (set of objects) is $X = \{x1, x2, x3, x4\}$, in which $x1$: hostile emotion; $x2$: obsessive-compulsive symptom; $x3$: interpersonal sensitivity; $x4$: paranoid symptoms. Among them, some of the inducement sets (attribute sets) for these psychological problems are $A = \{a1, a2, a3, a4, a5\}$, in which a1: professional attitude; $a2$: negative events; $a3$: economic problems; $a4$: ideal aspirations; $a5$: the only child. Through a certain amount of questionnaire test data collection, this paper establishes the relationship between object set X and attribute set A, such as table 1:

R	$a1$	$a2$	$a3$	$a4$	$a5$
$x1$	50.5	74.5	75.6	49.7	25.1
$x2$	14.3	85.7	85.6	42.9	14.3
$x3$	25.6	76.1	74.3	75.2	26.2
$x4$	40.2	81.2	60.3	58.9	21.2

Table 1. Proportion of Incentives for Psychological Problems (%)

According to the comprehensive analysis of the proportion of incentives and the actual investigation situation, the proportion of symptoms corresponding to incentives is high, the corresponding incentive membership degree is high, and the non-subordinate degree is low; The proportion of symptoms corresponding to incentives is low, the corresponding incentive non-subordinate degree is high, and the membership degree is low. This paper established the fuzzy relation $U(A \to X) = (uR(xi,ai), vR(xi,ai))$ between the set of objects and the attributes, in which $uR(xi, ai)$ refers to the membership function between two sets, and $vR(xi,ai)$ refers to the non-membership function between the two functions, according to the formula: lower approximation $\sum_{i=1}^{5} A^o(U) = (0.2, 0.7)$, upper approximation $\sum_{i=1}^{5} A^o(U) = (0.8, 0.1)$. For the counselling object, the symptom value of the psychological problem of the object is calculated as $h = (\mu', v')$, if it satisfies $\mu' > \mu$, $v' < v$, the user is deemed to have symptoms of some mental illness. Otherwise, the user is considered healthy. Compared with various psychological problems, diagnosis results are obtained: the higher the degree of agreement with the corresponding psychological problems, the more serious the corresponding psychological problems are.

3.2. Psychological Counseling Model

In a structured database of psychological problems, each of the trends in mental illness corresponds to a variety of counselling, but which one is the most appropriate needs a decision-making system to determine so as to improve the practicability of the counselling system. The intuitionistic fuzzy decision system $T = (U, C \cup D, V, f) A_i \subseteq C (i = 1, 2, 3, ..., m)$ is entered as a subset of condition attributes, $D = \{d\}$ denotes the decision attribute $X \in U / \{d\}$.

If $\sum_{i=1}^{m} R^o_{A_i}(x) \subseteq \sum_{k=1,k\neq j}^{m} R^o_{A_k}(x)$, Aj is relatively important in the condition attribute set, its importance is defined as:

$$sig(A_j, D) = \frac{\sum \left\{ \left| \sum_{k=1,k\neq j}^{m} R^o_{A_k}(x) \right| - \left| \sum_{i=1}^{m} R^o_{A_i}(x) \right| \right\}}{|U|} \tag{11}$$

If $\sum_{i=1}^{m} R^o_{A_i}(x) \subseteq \sum_{k=1,k\neq j}^{m} R^o_{A_k}(x)$, Aj is relatively unimportant in the conditional attribute set. In the consulting system, if sig$(Aj, D) = 0$, then the j^{th} psychological suggestion is not important in the whole decision system and can be neglected, if $sig(A_j, D) = \max_{k=1}^{m} sig(A_k, D)$, the j^{th} psychological advice is the most important in the entire decision-making system and must be prioritized.

4.Discussions

This article randomly selected 150 college students in a university to verify, with one of the students as an example: negative events and economic issues caused a lot of psychological pressure for him, and his professional attitude did not cause psychological pressure. So, it can be drawn that the disease value is:

A1 = (0.21,0.68)A2 = (0.82,0.10) A3 = (0.83,0.07) A4 = (0.45,0.48) A5 = (0.51,0.43) according to the formula,

$$\sum_{i=1}^{5} A^o(\mathrm{x}) = (0.21, 0.68), \sum_{i=1}^{5} A^o(x) = (0.83, 0.07) \tag{12}$$

So 0.2<0.21, 0.7>0.68, 0.8<0.83, 0.1>0.07, that is to determine that the classmates have some signs of mental illness. According to the formula:

$$\begin{aligned} &\sum_{i=1}^{5} A^o(x_1) = (0.2, 0.3), \overline{\sum_{i=1}^{5} A^o(x_1)} = (0.5, 0.2) \\ &\sum_{i=1}^{5} A^o(x_2) = (0.2, 0.7), \overline{\sum_{i=1}^{5} A^o(x_2)} = (0.8, 0.1) \\ &\sum_{i=1}^{5} A^o(x_3) = (0.3, 0.4), \overline{\sum_{i=1}^{5} A^o(x_3)} = (0.6, 0.1) \\ &\sum_{i=1}^{5} A^o(x_4) = (0.2, 0.4), \overline{\sum_{i=1}^{5} A^o(x_4)} = (0.7, 0.2) \end{aligned} \tag{13}$$

Therefore, psychological problems of the student are consistent with obsessive-compulsive symptoms, the diagnosis was more obvious obsessive-compulsive symptoms. For obsessive-compulsive symptoms, the psychological counseling plan x_j (j = 1, 2, 3, 4, 5) are, $x1$: arranging a reasonable schedule; $x2$: more sports; $x3$: intending to control their own behavior; $x4$: a useful psychological suggestion; $x5$: consult a psychiatrist. $a1$, $a2$, $a3$, $a4$, and $a5$ represent the respective indicator condition attributes of the object; $A1 = \{a1, a2\}, A2 = \{a3, a4\}, A3 = \{a5\}$ denote three decision indicators respectively, $\{d\}$ denotes the set of decision attributes. Intuitionistic fuzzy decision system T is input. Table 2 is an intuitive fuzzy decision-making system.

R	*a*1	*a*1	*a*3	*a*4	*a*5	*d*
x1	(0.4,0.5)	(0.3,0.5)	(0.8,0.2)	(0.4,0.5)	(0.7,0.1)	2
x2	(0.3,0.5)	(0.4,0.5)	(0.6,0.1)	(0.4,0.5)	(0.7,0.3)	2
x3	(0.3,0.5)	(0.4,0.4)	(0.4,0.5)	(0.4,0.5)	(0.7,0.3)	1
x4	(0.1,0.8)	(0.1,0.8)	(0.4,0.5)	(0.1,0.8)	(0.8,0.2)	1
x5	(0.7,0.3)	(0.4,0.5)	(0.9,0.1)	(0.5,0.5)	(0.8,0.1)	2

Table 2. Proportion of Incentives for Psychological Problems (%)

According to the formula, it can be calculated that $Sig\ (A1, D) = 0$, $Sig\ (A2, D) = 15$, $Sig\ (A3, D)=0$, so $A2$ is the most important relative to decision attributes.

5.Conclusion

In this special environment of university campus, the psychological condition of college students is worrying at present. How to detect psychological problems early and intervene has always been an important issue in college education and management. In this paper, aiming at the psychological health problems of undergraduates, based on the neural network algorithm, the prediction model of mental health of college students was proposed, and the factors related to depression and anxiety of college students were analyzed. The scientific and validity of the mental health prediction model were proved through the incomplete testing and verification of university undergraduates in a certain university. The model proposed in this paper can analyze the overall situation of college student's mental health and provide intelligent decision-making for the school students and management workers to do their mental health work well.

References

[1] Lejuez, C. W., Magidson, J. F., Mitchell, S. H., et al. (2010). Behavioral and biological indicators of impulsivity in the development of alcohol use, problems, and disorders. *Alcoholism: Clinical and Experimental Research*, 34(8), 1334-1345.

[2] Dierker, L., Stolar, M., Lloyd-Richardson, E., et al. (2008). Tobacco, alcohol, and marijuana use among first-year US college students: A time series analysis. *Substance Use & Misuse*, 43(5), 680-699.

[3] González, J. M. B., DesJardins, S. L. (2002). Artificial neural networks: A new approach to predicting application behavior. *Research in Higher Education*, 43(2), 235-258.

[4] Yan, W., Li, Y., Sui, N. (2014). The relationship between recent stressful life events, personality traits, perceived family functioning and internet addiction among college students. *Stress and Health*, 30(1), 3-11.

[5] Galbraith, C. S., Merrill, G. B., Kline, D. M. (2012). Are student evaluations of teaching effectiveness valid for measuring student learning outcomes in business-related classes? A neural network and Bayesian analyses. *Research in Higher Education*, 53(3), 353-374.

[6] Patock-Peckham, J. A., Cheong, J. W., Balhorn, M. E., et al. (2001). A social learning perspective: a model of parenting styles, self-regulation, perceived drinking control, and alcohol use and problems. *Alcoholism: Clinical and Experimental Research*, 25(9), 1284-1292.

[7] Oman, D., Shapiro, S. L., Thoresen, C. E., et al. (2008). Meditation lowers stress and supports forgiveness among college students: A randomized controlled trial. *Journal of American College Health*, 56(5), 569-578.