

# A confirmatory experimental study on the emotional Stroop effect in identifying depressive disorders among college students

Yuchen Chen<sup>1</sup>, Yanlin Sun<sup>2</sup>, Lifu Liang<sup>3\*</sup>

<sup>1</sup>Tianjin Juming Training School, Tianjin, China

<sup>2</sup>Tianjin University of Sport, Tianjin, China

<sup>3</sup>Tianjin Medical University, Tianjin, China

\*Corresponding Author. Email: 13114924460@163.com

---

**Abstract.** This study employed the emotional Stroop experiment from a behavioral research perspective to explore the role of the emotional Stroop effect in identifying depressive disorders among college students and to examine its correlation with the results of the Self-Rating Depression Scale (SDS) and the Patient Health Questionnaire-9 (PHQ-9). The results showed that the emotional Stroop effect of depressive words was significantly correlated with the scores of the depression scales, with correlations of  $r = 0.757$  ( $p < 0.01$ ) for the SDS and  $r = 0.739$  ( $p < 0.01$ ) for the PHQ-9. These findings confirm the validity of the emotional Stroop effect in identifying depressive disorders among college students, indicating that it serves as an effective supplementary method to depression scale assessments for screening depressive disorders.

**Keywords:** emotional Stroop effect, college students, depressive disorder, identification

---

## 1. Introduction

Depressive disorder refers to a type of mood disorder characterized primarily by prominent and persistent depressive symptoms caused by various factors. Its main features include low mood, slowed thinking, diminished interest, and weakened volition, often accompanied by other psychological and somatic symptoms [1]. In recent years, college students have faced increasing pressure from multiple sources, such as academics, employment, interpersonal relationships, and emotions, leading to a steady rise in the incidence of depressive disorders within this population. Depression has become one of the most common psychological disorders among college students and is a leading cause of suicide and academic withdrawal in universities. Epidemiological data indicate that the incidence rate of depressive disorders among college students has reached 24%, with depressive tendencies and dysthymia being the main manifestations. This suggests that college students have become a high-risk group for depressive disorders [2-4]. Therefore, early identification and intervention of depression in college students are of great significance for reducing depressive symptoms and improving overall mental health. This study adopts the emotional Stroop experimental paradigm—a method widely used in psychology to study the relationship between emotion and cognition—to explore its effectiveness in identifying depressive disorders among college students.

## 2. Participants and methods

### 2.1. Participants

A total of 153 college student volunteers were recruited from a university in Tianjin, including 66 males and 87 females. Among them, 32 were graduate students (27 master's and 5 doctoral) and 121 were undergraduates. The average age was  $22.41 \pm 3.65$  years. All participants had normal or corrected-to-normal vision, no color blindness or color weakness, and were all right-handed.

### 2.2. Experimental design

This study adopted a single-factor repeated measures design. The independent variable was the type of word (positive, neutral, depressive), and the dependent variables were reaction time, number of errors, and number of omissions.

## 2.3. Materials

### 2.3.1. Emotional words

Thirty depressive adjectives were randomly selected from subscales E, F, and G of the Depression Adjective Checklist [1], and thirty neutral adjectives were randomly selected from the Modern Chinese Dictionary. Eight graduate students majoring in psychology rated the emotional valence and familiarity of the sixty adjectives (30 depressive and 30 neutral) using a 5-point scale. The scale ratings were as follows: 1 = depressive, 2 = slightly depressive, 3 = neutral, 4 = slightly positive, 5 = positive. The evaluation results are shown in Table 1.

**Table 1.** Evaluation Results of Emotional Valence and Familiarity of Words ( $n = 60$ )

	<i>n</i>	Emotional Valence		Familiarity	
		Mean Valence	<i>SD</i>	Mean Familiarity	<i>SD</i>
Depressive Words	30	1.375	0.553	4.256	0.312
Neutral Words	30	3.152	0.434	4.417	0.256

A valence score of 3 was used as the criterion for statistical testing. Depressive words were selected if their emotional valence was significantly lower than 3, while neutral words were selected if their emotional valence did not significantly differ from 3. All words chosen had a familiarity rating higher than 3. Ultimately, 20 depressive adjectives and 20 neutral adjectives that met the criteria were selected as experimental materials. The students who participated in material evaluation did not take part in the formal experiment.

### 2.3.2. Scales

#### 2.3.2.1. Zung Self-Rating Depression Scale (SDS)

The Zung Self-Rating Depression Scale (SDS), developed under the direction of Zung, is a self-report tool designed to measure the severity of depressive symptoms. The scale consists of 20 items, primarily assessing symptoms related to depressive states, and provides an intuitive reflection of the respondent's subjective feelings. The split-half reliability of the scale was reported as 0.92 (1986), and the criterion-related validity (with HRSD) was 0.84 [5].

#### 2.3.2.2. Patient Health Questionnaire-9 (PHQ-9)

The Patient Health Questionnaire-9 (PHQ-9) is a 9-item self-assessment instrument for screening depressive symptoms, based on the diagnostic criteria for depression in the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). It has a Cronbach's  $\alpha$  coefficient of 0.85, a 4-week test-retest reliability of 0.88, and structural validity ranging from 0.618 to 0.761 [6].

## 2.4. Procedure

### 2.4.1. Emotional stroop experiment

The experiment was programmed and conducted using the Display Master DX (DMDX) psychological experiment software system. Stimuli were presented via computer, and participants' responses were recorded automatically. The experiment consisted of two parts: a preparatory practice session followed by the formal test.

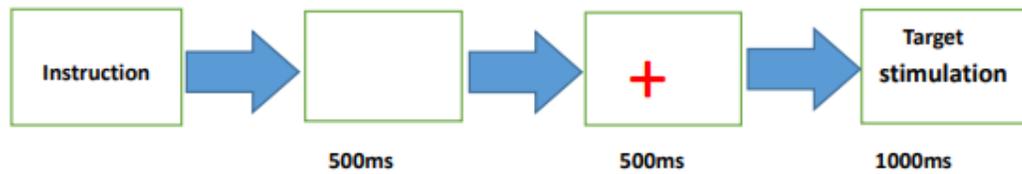
#### 2.4.1.1. Preparatory practice

To minimize experimental errors due to unfamiliarity with color or keyboard layout, participants first completed a preparatory practice session. In this stage, four colored shapes (red, green, yellow, and blue) were randomly presented on the computer screen. Participants were instructed to respond using corresponding keyboard keys: A -Red, D-Green, J-Yellow, L-Blue. Only participants who achieved 100% accuracy during practice proceeded to the formal test.

#### 2.4.1.2. Formal experiment

The experimental stimuli were presented in random order. DMDX recorded each participant's reaction time, number of omissions, and number of errors. Experimental Procedure: Instructions → Blank Screen (500 ms) → Fixation Point (500 ms) →

Target Stimulus (1,000 ms). The target stimulus disappeared immediately after the participant's response; if no response was made, it disappeared automatically after 1,000 ms (see Figure 1).



**Figure 1.** Procedure of the Emotional Stroop Experiment

## 2.5. SDS and PHQ-9 tests

After completing the emotional Stroop experiment, 37 participants who exhibited an emotional Stroop effect were included in the study group. These participants subsequently completed the SDS and PHQ-9 depression scales.

## 2.6. Data processing

After the experiment, all participant data were compiled. Data from participants with an error rate greater than 5% or who failed to follow instructions were excluded. All valid data were statistically analyzed using SPSS 23.0 software.

# 3. Results

## 3.1. Results of the emotional stroop experiment among college students

The results of reaction time, number of errors, and omissions for depressive and neutral words in the emotional Stroop experiment are presented in Table 2.

**Table 2.** Reaction Time, Number of Omissions, and Number of Errors in the Emotional Stroop Experiment ( $M \pm SD$ )

Reaction Type	Depressive Words	Neutral Words	<i>t</i>
Reaction Time	557 ± 4.4410	522 ± 2.5166	41.708**
Number of Errors	1.189 ± 1.0498	0.622 ± 0.7208	2.711**
Number of Omissions	0.676 ± 0.7837	0.243 ± 0.4350	2.935**

Note: \* $p < 0.05$ , \*\* $p < 0.01$

The results in Table 2 indicate that in the emotional Stroop experiment, the reaction time for depressive words was significantly longer than that for neutral words ( $t = 41.708, p < 0.01$ ), demonstrating interference of emotional information with non-emotional information processing—namely, the emotional Stroop effect. Furthermore, significant differences were also found in omission rate ( $t = 2.935, p < 0.01$ ) and error rate ( $t = 2.711, p < 0.01$ ) between depressive and neutral words, suggesting that the participants exhibited deficiencies in distraction inhibition and attentional bias, consistent with previous findings [7, 8]. Since the overall error and omission rates of both groups were below 0.5%, these differences were considered negligible and were excluded from further statistical analysis.

## 3.2. Results of the SDS and PHQ-9 depression scale tests

After completing the emotional Stroop experiment, the participants in the study group were assessed using the SDS and PHQ-9 depression scales. The results are presented in Table 3.

**Table 3.** Results of the SDS and PHQ-9 Depression Scales among College Students ( $M \pm SD$ )

Group	SDS	PHQ-9
Study Group	59.89 ± 6.691	8.54 ± 5.747

Table 3 shows that the mean SDS score was  $59.89 \pm 6.691$ , and the mean PHQ-9 score was  $8.54 \pm 5.747$ . The distribution of participants across different severity levels is shown in Table 4 and Table 5.

**Table 4.** Distribution of Participants by SDS Depression Severity Levels

Distribution	53 ≤ No Depression	53-62 Mild Depression	63-72 Moderate Depression	≥ 73 Severe Depression
<i>n</i>	7	18	10	2
%	18.92%	48.65%	27.03%	5.41%

The SDS results were scored using raw scores according to the manual's guidelines:  $\leq 53$  = no depression, 53-63 = mild depression, 63-72 = moderate depression, and  $\geq 73$  = severe depression. As shown in Table 3 and Table 4, 7 participants (18.92%) showed no depression, 18 (48.65%) showed mild depression, 10 (27.03%) showed moderate depression, and 2 (5.41%) showed severe depression. In total, 30 participants (81.08%) fell within the mild-to-severe depression range. These findings indicate that among individuals who exhibited an emotional Stroop effect, the positive detection rate of depression using the SDS scale was 81.08%.

**Table 5.** Distribution of Participants by PHQ-9 Depression Severity Levels

Distribution	0-4 No Depression	5-9 Mild Depression	10-14 Moderate Depression	15-19 Moderately Severe Depression	20-27 Severe Depression
<i>n</i>	8	17	8	2	2
%	21.62%	45.95%	21.62%	5.41%	5.41%

According to the PHQ-9 scoring manual, a total score of 0-4 indicates no depression, 5-9 mild depression, 10-14 moderate depression, 15-19 moderately severe depression, and 20-27 severe depression. As shown in Tables 3 to Table 5, 8 participants (21.62%) had no depression, 17 (45.95%) had mild depression, 8 (21.62%) had moderate depression, 2 (5.41%) had moderately severe depression, and 2 (5.41%) had severe depression. Altogether, 29 participants (78.38%) fell within the mild-to-severe depression range. These findings suggest that among individuals who exhibited an emotional Stroop effect, the positive detection rate of depression using the PHQ-9 scale was 78.38%.

### 3.3. Correlation between reaction time in the emotional stroop experiment and depression scale scores (SDS and PHQ-9)

To further verify the effectiveness of the emotional Stroop effect in identifying depressive disorders among college students, correlation analyses were conducted between reaction times in the Stroop task and the scores of the SDS and PHQ-9 scales. The results are shown in Table 6.

**Table 6.** Correlations Between the Emotional Stroop Effect and Depression Scale Scores (SDS and PHQ-9)

Word Type	SDS	PHQ-9
Neutral Words	0.091	0.073
Depressive Words	0.757**	0.739**

Note: \* $p < 0.05$ , \*\* $p < 0.01$

As shown in Table 6, the reaction time for depressive words was significantly correlated with the SDS score ( $r = 0.757$ ,  $p < 0.01$ ), while the reaction time for neutral words showed no significant correlation ( $r = 0.091$ ,  $p > 0.05$ ). Similarly, the reaction time for depressive words was significantly correlated with the PHQ-9 score ( $r = 0.739$ ,  $p < 0.01$ ), whereas the reaction time for neutral words was not ( $r = 0.073$ ,  $p > 0.05$ ). These results demonstrate that when SDS and PHQ-9 are used as criterion measures, the emotional Stroop effect shows a high level of criterion-related validity with both scales, further confirming that the emotional Stroop effect is an effective tool for identifying depressive disorders among college students.

## 4. Discussion

This study conducted a confirmatory experiment to examine the effectiveness of the emotional Stroop effect in identifying depressive disorders among college students. The emotional Stroop effect is an experimental paradigm used to investigate how emotional information interferes with the cognitive processing of non-emotional information [9]. By engaging both emotional

and cognitive processing through a dual-task parallel structure, the emotional Stroop paradigm provides insights into attentional bias and cognitive processing mechanisms in response to emotional stimuli [10]. It is a classic method for exploring the interaction between emotion and cognition.

In this experiment, the emotional Stroop task was first used to screen 37 participants who demonstrated an emotional Stroop effect from the total pool of recruited college students. These 37 participants were then assessed using the Self-Rating Depression Scale (SDS) and the Patient Health Questionnaire-9 (PHQ-9) to determine the prevalence of depressive disorders and to verify the validity of the emotional Stroop effect in identifying such conditions.

The results showed that among the 37 participants with an emotional Stroop effect, the SDS detected 18 with mild depression, 10 with moderate depression, 2 with severe depression, and 7 without depression. Combining mild to severe cases yielded 30 individuals, accounting for 81.08% of the total, indicating an overall depression detection rate of 81.08% by the SDS among participants showing the emotional Stroop effect. Similarly, the PHQ-9 identified 17 with mild depression, 8 with moderate depression, 2 with moderate-to-severe depression, 2 with severe depression, and 8 without depression, resulting in an overall detection rate of 78.38%.

To further validate the effectiveness of the emotional Stroop effect in identifying depressive disorders among college students, correlation analyses were conducted between the emotional Stroop results and those of the SDS and PHQ-9. The findings revealed a significant correlation between the emotional Stroop effect and the SDS results ( $r = 0.757, p < 0.01$ ), as well as with the PHQ-9 results ( $r = 0.739, p < 0.01$ ). These results confirm that the emotional Stroop effect has significant criterion-related validity in identifying depressive disorders among college students.

In addition, the secondary screening results from the SDS and PHQ-9 indicated that mild and moderate depression were the predominant forms among the students. The SDS identified 28 participants (75.68%) with mild or moderate depression, while the PHQ-9 identified 25 participants (67.57%) in these categories—both exceeding half of the total number of students identified with depression by the Stroop experiment. Only two participants (5.41%) were classified as having severe or moderate-to-severe depression. This aligns with observations from practical counseling, where severe depressive disorders are relatively rare among college students. Moreover, results from both the SDS and PHQ-9 indicated that depressive disorders among college students are primarily characterized by depressive mood, somatic symptoms (such as sleep disturbances, appetite loss, constipation, and fatigue), and cognitive impairments (including poor concentration, memory decline, and negative thinking). A few students with severe depression exhibited suicidal ideation or tendencies. These findings highlight the importance of university mental health education and counseling programs, especially for students showing signs of suicidal thoughts or behaviors.

This study further confirms that college students with depressive emotions tend to exhibit cognitive structures biased toward negative evaluation and interpretation of information, and they preferentially process information congruent with their emotional state [11-14]. This insight provides valuable implications for psychological interventions in depressive disorders.

It is also noteworthy that although some students' SDS and PHQ-9 scores fell below the clinical cutoff for depression, several of them displayed partial positive symptoms. These individuals may fall within the subthreshold depression range and could potentially develop full depressive disorders if not given timely intervention. Therefore, special attention should be given to this group in university mental health education and counseling.

## 5. Conclusion

This study verified the effectiveness of the emotional Stroop effect in identifying depressive disorders among college students. The results demonstrate that the emotional Stroop effect is an effective method for identifying depressive disorders and serves as a valuable supplement to traditional scale-based screening approaches.

## References

- [1] National Health Commission Medical Administration and Medical Reform Bureau. (2020). *Standards for the Diagnosis and Treatment of Mental Disorders*. Beijing: People's Medical Publishing House.
- [2] Du, Z. Y., & Wang, K. Q. (1999). Epidemiological survey on depression among 1, 597 college students. *Chinese Journal of Behavioral Medicine Science*, 8(3), 172–173.
- [3] Li, T. (2008). A study on depression and its related factors among college students. *Social Psychological Science*, 23(100), 547–553.
- [4] Cui, Q. X., & Wang, Z. X. (2014). Investigation and analysis on the status and influencing factors of depression among college students. *China Health Administration*, (8), 629–630+633.
- [5] Wang, X. D. (1999). Manual for psychological health scale assessment (revised edition). *Chinese Journal of Mental Health (Suppl.)*, 194–200.
- [6] Li, L. J., & Ma, X. (Eds.). (2015). *Guidelines for the prevention and treatment of depressive disorders in China* (2nd ed.). Chinese Medical Multimedia Press.
- [7] Chen, Y. C. (2022). *The role of the Emotional Stroop effect in the identification of depression disorders among college students* [Master's thesis, Tianjin University of Sport].

- [8] Liang, L. F. (2009). An experimental study on the emotional Stroop effect among college students. *Journal of Tianjin Medical University*, 15(3), 512–514.
- [9] Frank, P., McKenna, D., & Dinkar, S. (1995). Intrusive cognitions: An investigation of the emotional Stroop task. *Journal of Experimental Psychology*, 21(6), 1595–1607.
- [10] MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology*, 95(1), 15–20.
- [11] Lichtenstein-Vidne, L., Okon-Singer, H., Cohen, N., Mertens, J., & Henik, A. (2017). Attentional bias in clinical depression and anxiety: The impact of emotional and non-emotional distracting information. *Biological Psychology*, 122, 4–12.
- [12] Bradley, B. P., Mogg, K., & Lee, S. C. (1997). Attentional biases toward negative information in induced and naturally occurring dysphoria. *Behaviour Research and Therapy*, 35(16), 911–927.
- [13] Platt, B., Murphy, S. E., & Lau, J. Y. (2015). The association between negative attention biases and symptoms of depression in a community sample of adolescents. *PeerJ*, 3, e1372.
- [14] Fox, E., & Dutton, K. (2002). Attentional bias for threat: Evidence for delayed disengagement from emotional faces. *Cognition and Emotion*, 16(3), 355–379.