



Single online self-compassion writing intervention reduces anxiety: With the feedback of ChatGPT

Tomoko Kishimoto ^{a,b,*¹}, Ximing Hao ^{b,1}, Trimble Chang ^{b,c}, Zhiye Luo ^b

^a Faculty of Psychology, Beijing Normal University, Beijing, China

^b Department of Social Psychology, School of Sociology, Nankai University, Tianjin, China

^c Department of Intelligent Science, College of Artificial Intelligence, Nankai University, Tianjin, China



ARTICLE INFO

Keywords:

Self-compassion writing
Online intervention
Single-session intervention
Anxiety
ChatGPT feedback

ABSTRACT

With the aim of universal prevention, there is still a lack of non-clinician-led or self-help interventions designed for the general population experiencing psychological distress. This study aims to address this gap by exploring a convenient, single-session, and effective method for reducing anxiety using ChatGPT feedback. Using a mixed-factorial experimental design, we recruited 98 participants online to compare changes in self-compassion and anxiety before and after the intervention among three groups: self-compassion writing with ChatGPT feedback (Experimental Group 1), self-compassion writing alone (Experimental Group 2), and an Active Control Group. Results showed that both Experimental Group 1 and 2 experienced significant increases in self-compassion and Experimental Group 1 experienced significant decreases in anxiety after the intervention, with Experimental Group 1 showing greater improvements. There were no significant changes in the control group. This study demonstrates that single online self-compassion writing intervention can effectively reduce anxiety and that using ChatGPT to provide highly empathetic feedback can enhance the effectiveness, making psychological support more accessible, cost-effective, and suitable for the general population.

1. Introduction

As an emotional disorder that most people experience in their lifetime, anxiety is a future-oriented mood state characterized by worry about potential negative events, threat-biased appraisals, and accompanied by behaviors like avoidance and physiological responses such as muscle tension and autonomic activation (Craske et al., 2011). Social competition, survival pressures, and the COVID-19 pandemic have exacerbated the spread of anxiety, making it a significant issue affecting people's mental health (Mazza et al., 2020). A meta-analysis that includes data from seven countries (including China) shows that the proportion feeling distressed by anxiety surged to 38.12 % during the pandemic (Necho et al., 2021). Prolonged and unresolved anxiety can lead to various physiological and psychological issues, such as mood disorders (Brown et al., 2001), sleep disorders (Cox and Olatunji, 2016), digestive problems and decreased immune system function (Januzzi et al., 2000), and may also affect cognitive function (Leonard and Abramovitch, 2019), social relationships and work performance (McCarthy et al., 2016). According to Bandelow and Michaelis (2015),

only 20.6 % of individuals with anxiety disorders worldwide seek medical help, leaving the remaining population at risk of persistent or worsening symptoms that are difficult to self-heal. Therefore, timely self-regulation or intervention measures to prevent the excessive spread of anxiety are crucial (Powell and Enright, 2015). Self-compassion writing, as a method of emotional regulation, has been shown to produce positive changes in psychological (depression, anxiety, positive/negative affect), and physical (pain) variables (Aydoğdu and Dirik, 2022).

Research has shown that offline self-compassion writing interventions can effectively alleviate stress levels, reduce anxiety, and enhance psychological well-being. For example, Wong and Mak (2016) conducted a week-long self-compassion writing intervention among students in Hong Kong, and found that the experimental group exhibited better physical health compared to the control group. Odou and Brinker (2015) found that by engaging in self-compassion writing about past distressing or stressful events, individuals can quickly grasp the core principles of self-compassion, promote positive coping with stress, and reduce related psychological symptoms. However, offline self-

* Corresponding author at: No. 19, XinJieKouWai St., HaiDian District, Beijing 100875, China.

E-mail address: tomokopp@126.com (T. Kishimoto).

¹ Tomoko Kishimoto and Ximing Hao have contributed equally as co first authors.

compassion writing interventions vary in frequency from single to multiple sessions, rely on professional guidance, and the offline mode limits large-scale implementation, making it difficult to meet the universal prevention needs of the general population.

A single-session online self-compassion writing intervention may address this issue, which have been proposed to increase accessibility and reduce the financial constraints that typically accompany longer-term treatments (Schleider and Weisz, 2017). Although most online self-compassion writing interventions currently involve multiple sessions, research has demonstrated the effectiveness of single-session online self-compassion writing interventions. For instance, Guan et al. (2021) conducted an online self-compassion writing guided by therapists during the COVID-19 pandemic and found it effectively reduced participants' anxiety and alleviated negative emotions. However, these interventions typically require guidance from therapists or other professionals, which limits their potential as universal prevention measures for the general population. To address this issue, it is crucial to find new methods to remove the dependence on human therapists while maintaining efficacy, thus reducing the costs of these interventions.

As a large language model developed by OpenAI, ChatGPT has the potential to serve as an assistant to therapists, or even replace therapists in certain interventions in the field of psychotherapy. For example, Carlbring et al. (2023) studied ChatGPT as a virtual mentor for self-help interventions, interacting with patients, providing analysis and feedback. Nazarova (2023) explored ChatGPT as an AI therapist, identifying patients' cognitive distortions and posing thought-provoking questions to challenge negative thinking. Existing application research indicates that using ChatGPT in clinical psychotherapy can enhance the effectiveness of online self-help interventions and reduce costs. In this context, we aim to explore the feasibility of using ChatGPT to provide feedback for online self-help interventions, with the goal of achieving a more cost-effective and efficient anxiety intervention model.

Based on the above review, there remains a lack of self-help interventions designed for the general population experiencing psychological distress aimed at universal prevention. Implementing interventions using ChatGPT to provide feedback for online self-compassion writing interventions can improve effectiveness and accessibility. In this study, we want to explore the effectiveness of an online self-compassion writing intervention, with ChatGPT providing feedback. By providing participants with timely, empathetic feedback, it is expected to effectively alleviate anxiety in the general population, reduce intervention costs, increase accessibility, and promote the rational distribution of mental health service resources (Kazdin, 2019). Additionally, we aim to fill the gap in online self-help interventions designed for universal prevention and targeting the general population.

We want to employ a mixed-factorial experimental design through an online self-compassion writing intervention to increase self-compassion and positive emotions, reduce anxiety and negative emotions. The experiment will compare the differences in variables before and after the intervention among three groups: self-compassion writing with ChatGPT feedback (Experimental Group 1), self-compassion writing alone (Experimental Group 2), and Active Control Group. In sum, we test the following hypotheses:

Hypothesis 1. Online self-compassion writing intervention will significantly increase self-compassion and positive affect and significantly reduce anxiety and negative affect.

Hypothesis 2. Compared to self-compassion writing alone, participants with the feedback of ChatGPT will achieve better outcomes, showing significant increases in self-compassion and positive affect, and reductions in anxiety and negative affect.

2. Methods

Study design, hypotheses, and analysis plan were pre-registered with the Open Science Framework (see doi:[10.17605/OSF.IO/E5S9Z](https://doi.org/10.17605/OSF.IO/E5S9Z)).

2.1. Participants and recruitment

We referenced the self-compassion writing intervention data from Mei et al. (2023) and used G*Power 3.1 to calculate the required sample size. The parameters set were: ANOVA (Repeated measures, within-between interaction), effect size $f = 0.25$, power $1 - \beta = 0.95$, $\alpha = 0.05$, number of groups = 3, number of measurements = 3, resulting in a minimum required sample size of 54 participants.

Since our research aims to provide preventive interventions for a broad population experiencing emotional issues, we did not need to specifically recruit clinical patients with mental disorders. Participants only need to be at least 18 years old, with no special requirements regarding demographic characteristics or clinical psychological indicators.

Participants were recruited through online platforms and placed into three groups above in order of recruitment. The study was conducted in three phases, each phase implementing the intervention with one group. We included data from participants who authorized access to the Mini Program and completed all intervention tasks, resulting in a final sample of 35 participants in Experimental Group 1, 33 in Experimental Group 2, and 30 in the Active Control Group. Thus, the final sample consisted of 98 participants ($n = 64$ female, 65.3%; $n = 98$ yellow, 100%).

2.2. Assessment

2.2.1. Self-Compassion Scale (SCS)

Developed by Neff (2003), the SCS is the most commonly used tool to measure self-compassion levels. It consists of 26 items, using a 1–5 point rating scale. The average score of all items represents the overall level of self-compassion, with higher scores indicating higher self-compassion. The Chinese version of the SCS has demonstrated good reliability and validity (Chen et al., 2011), with a Cronbach's α of 0.92 in this study.

2.2.2. State-Trait Anxiety Inventory (STAI)

Developed by Spielberger et al. (1970), the STAI is used to assess state anxiety under stress and trait anxiety as a personality characteristic. It comprises 40 items divided into two subscales: state anxiety (items 1–20) and trait anxiety (items 21–40). Items are rated on a 1–4 point scale, with higher scores indicating higher anxiety. Since state anxiety can be alleviated through short-term interventions, whereas trait anxiety is more stable and difficult to change in the short term (Spielberger et al., 1970), we only used items 1–20 to better evaluate the effectiveness. The Chinese version of the STAI has shown good reliability and validity (Li and Qian, 1995), with a Cronbach's α of 0.91 for items 1–20 in this study.

2.2.3. Positive and Negative Affect Schedule (PANAS)

Developed by Watson et al. (1988), the PANAS measures immediate positive affect (PA) and negative affect (NA). It consists of 20 items divided into two subscales: PA and NA, using a 1–5 point rating scale. Higher scores indicate stronger emotions. The Chinese version of the PANAS has demonstrated good reliability and validity (Huang et al., 2003), with a Cronbach's α of 0.88 for PA and 0.85 for NA in this study.

2.3. Experiment procedures

WeChat is the most widely used social platform in China (Montag et al., 2018), making it a convenient tool for participant engagement. So we conducted this study through a self-developed WeChat Mini Program, which enabled online answering and ChatGPT's feedback. At the beginning of the experiment, participants in all three groups completed the Self-Compassion Scale (SCS; Chen et al., 2011), the State-Trait Anxiety Inventory (STAI; Li and Qian, 1995) questions 1–20, and the Positive and Negative Affect Scale (PANAS; Huang et al., 2003).

After completing the scales, all participants began a negative event writing task. Following this, all participants filled out the PANAS to

assess the task effects. Then Experimental Group 1 and 2 engaged in the self-compassion writing task. In this task, Experimental Group 1 received empathetic feedback generated by ChatGPT based on the core elements of self-compassion and the content they wrote. Experimental Group 2 did not receive any feedback. Meanwhile, Active Control Group performed a neutral writing task.

After completing the self-compassion writing or the neutral writing task, all participants again filled out the SCS, STAI (questions 1–20), and PANAS to assess the intervention effects. The entire experiment took approximately 30 min to complete.

The study design is shown in Fig. 1.

2.4. Interventions

2.4.1. Negative event writing task

To activate negative emotions and prepare them for the subsequent self-compassion writing task, participants were required to describe in detail a negative event from their past that made them feel terrible (see Fig. 2). The writing duration was 10 min. Previous research indicates that writing about negative events typically causes immediate, temporary emotional distress and short-term discomfort (Baikie and Wilhelm, 2005).

2.4.2. Self-compassion writing task

Participants in Experimental Group 1 and 2 should complete the self-compassion writing task. In this task, participants wrote based on prompts for a total duration of 10 min. The first prompt induced a sense of common humanity, asking them to list ways in which other people might experience similar negative events. The second prompt induced self-kindness, requiring them to express understanding, kindness, and care toward themselves as they would toward a close friend experiencing a similar event. The final prompt induced mindfulness, asking them to describe their feelings about the event in an objective, non-emotional manner (Leary et al., 2007) (see Fig. 2). ChatGPT provided empathetic feedback based on these elements of self-compassion. The basic intervention model and examples of ChatGPT's involvement are shown in Fig. 2.

2.4.3. Neutral writing task

To validate the effectiveness of the self-compassion writing task, we need to balance the writing time and the amount of content for the Active Control Group, while also managing their negative emotions. Following the study by Wong and Mak (2016), the Active Control Group

was asked to describe their daily activities from yesterday, today, and tomorrow in a factual and unemotional manner. The writing duration was also 10 min.

2.4.4. ChatGPT training

For Experimental Group 1, the Mini Program was designed to integrate ChatGPT by inserting specific code to interface with the ChatGPT 4.0 API. This integration allowed ChatGPT to provide real-time feedback to participants upon completing the self-compassion writing task.

The training prompts for ChatGPT's feedback were created by us (see Fig. 2). Feedback focused on the three elements of self-compassion tailored to the negative events reported by participants, and provided highly empathetic support and advice. We aimed to use this empathy technique to enhance self-compassion and reduce anxiety during the interventions.

In the figure, *italic font* represents the content entered by participants, others are the instructions or feedback presented to participants during the intervention. Solid boxes represent the actual intervention process, while dashed box represents the backend processing. The example in the figure is sourced from an actual participant response during the experiment, with private information anonymized. We have translated it from Chinese into English.

2.5. Data analysis

We used SPSS 29.0 and RStudio for data analysis.

2.5.1. Propensity score matching

Due to practical constraints during the study's implementation, the experiment was not conducted as a randomized controlled trial (RCT). Instead, three groups were conducted sequentially. To ensure that the data better reflect an RCT setting and to minimize the confounding effect of the experiment's timing, we applied Propensity Score Matching (PSM).

We used one-to-one matching, with the Active Control Group as the reference. Experimental Group 1 and 2 were matched to the Active Control Group. After matching, each of three groups had 30 participants remaining. To assess the effectiveness of the matching, we referred to the Standardized Mean Difference (SMD) metric. Generally, an SMD absolute value of less than 0.25 indicates a good matching effect (Lutz et al., 2016).

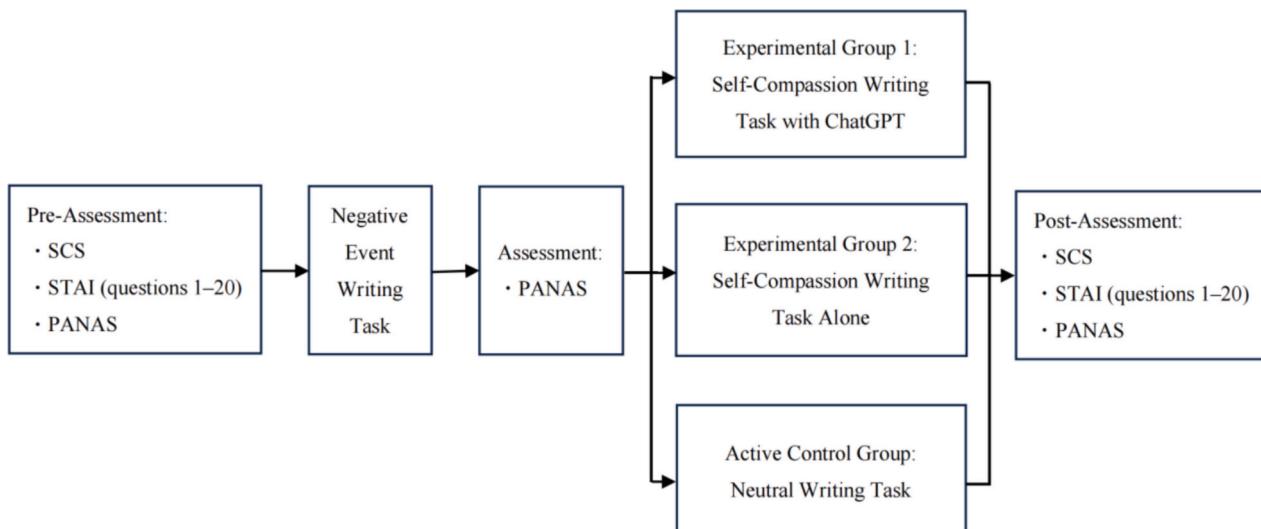


Fig. 1. Study design.



Fig. 2. Training and Feedback Principles, Basic Model and Example of ChatGPT's Involvement.

2.5.2. Repeated measures ANOVA

We conducted a repeated measures ANOVA to test the hypotheses and explore the effects of online self-compassion writing and the feedback of ChatGPT on self-compassion and anxiety, as well as the impact of the negative event writing, self-compassion writing, and neutral writing task on emotional states. Within-subjects factor was time, and between-subjects factor was group. The repeated measures ANOVA was performed on self-compassion, anxiety, and positive and negative affect. After completing the analysis, we reported the main effects of time and group, as well as their interaction effects. For the significant main effects, we conducted simple effects analyses; and for significant interaction effects, we performed post hoc multiple comparisons.

2.5.3. Natural language processing

Since we conducted this study online, natural language processing (NLP) of participants' writing content was more convenient compared to offline experiments. [Vine et al. \(2020\)](#) demonstrated that the richness of emotional vocabulary (EVs) is correlated with typical emotional experiences. The formulas for EVs is as follows.

The Richness of Emotional Vocabulary (EVs)

$$= \left(\frac{\text{Number of Emotional Words excluding repetitions}}{\text{Total Word Count}} \right) \times 100$$

Meanwhile, one of the core elements of self-compassion is mindfulness ([Neff, 2003](#)), which shows the ability to perceive and understand emotional experiences when facing negative events. Based on these conclusions, we used the total word count and EVs obtained through NLP as indirect measures to assess the differences between each groups. The total word count can reflect the participants' ability to express themselves and the seriousness of their answers. EVs can indirectly reflect the ability to perceive and express emotions, which served to validate the effectiveness of the intervention.

For NLP, we first used the Chinese psychological text analysis system developed by [Gao et al. \(2013\)](#) to segment the writing texts and obtain the total word count. Next, we used custom R scripts to calculate the EVs. These scripts referenced the HowNet Sentiment Dictionary ([Dong et al., 2010](#)), the NTUSD Simplified Chinese Sentiment Dictionary from National Taiwan University ([Ku et al., 2006](#)), and Tsinghua University's Li Jun Chinese Polarity Lexicon ([Li, 2011](#)) to extract emotional words and calculate their numbers.

3. Results

3.1. Propensity score matching

The Propensity Score Matching (PSM) results (Standardized Mean Differences, SMD) for Experimental Group vs. Active Control Group are shown in [Table 1](#). As observed, aside from the less-than-ideal matching effect between Experimental Group 1 and Active Control Group for positive affect, the matching effect for all other groups was satisfactory. One-way ANOVA was conducted on the baseline of positive affect, and results showed that there was no difference among the three groups, $F(1, 87) = 2.30, p = .106$. Therefore, the data from this study can approximately be analyzed as randomized controlled trial data.

Table 1
The PSM results for Experimental Group vs. Active Control Group (SMD).

Variables	Exp. 1. vs. Act.		Exp. 2. vs. Act.	
	Before	After	Before	After
Self-Compassion	5.80 %	6.42 %	6.82 %	-3.14 %
Anxiety	-12.03 %	-9.93 %	-10.32 %	-5.86 %
Positive Emotion	-43.06 %	-52.55 %	-11.93 %	-12.23 %
Negative Emotion	1.31 %	2.84 %	-19.56 %	-13.74 %

Notes: Exp. means Experimental Group, Act. means Active Control Group, same as below.

3.2. Manipulation test: total word count and EVs

Since ChatGPT intervened only after the completion of the writing tasks, it did not directly influence the writing texts. Therefore, Experimental Group 1 and 2 were combined for analysis. NLP results are shown in [Table 2](#).

In terms of total word count, there was no significant difference between Experimental Group and Active Control Group in both types of tasks, indicating that the participants' expression ability and the seriousness of their answers were approximately the same. In the negative event writing task, there were no significant differences between the two groups in the EVs, suggesting that participants in both groups performed similarly in the negative event writing task. However, Experimental Group showed significantly higher EVs compared to Active Control Group in the writing intervention task, which indicates that self-compassion writing task indeed enhanced the intensity of participants' typical emotional experiences related to positive and negative emotions, confirming the effectiveness of the experimental manipulation.

3.3. Demographic variables and descriptive statistics

Due to considerations of anonymity, we collected minimal precise private information from participants. Age was collected in broad intervals to maintain privacy. The gender and age distribution of participants and the descriptive statistics are shown in [Table 3](#). χ^2 analyses were conducted for each demographic variable, revealing no significant differences in gender and age among three groups. One-way ANOVA was conducted for each variable, indicating no significant differences among three groups except the post-assessment of anxiety.

3.4. Main results

Repeated measures ANOVA was conducted on self-compassion and anxiety, with time (pre, post) as the within-subjects factor and group (Experimental Group 1, 2, and Active Control Group) as the between-subjects factor.

3.4.1. Self-compassion

The main effect of time was significant, $F(1, 87) = 21.66, p < .001$, $\eta_p^2 = 0.199$. The main effect of group was not significant. The interaction effect between time and group was significant (see [Fig. 3](#)), $F(2, 87) = 7.01, p = .002, \eta_p^2 = 0.139$.

Simple effects analysis showed that self-compassion significantly increased after the intervention in Experimental Group 1, $F(1, 87) = 29.96, p < .001, \eta_p^2 = 0.250$; and 2, $F(1, 87) = 6.70, p = .011, \eta_p^2 = 0.072$. The Active Control Group showed no significant change. Additionally, post-intervention self-compassion in Experimental Group 1 were significantly higher than those in Active Control Group ($p = .034$). There was no significant difference between Experimental Group 1 and 2, or between Experimental Group 2 and Active Control Group.

Results indicate that self-compassion writing intervention significantly increases participants' self-compassion, regardless of ChatGPT involvement. However, the involvement of ChatGPT enhances the effectiveness in increasing self-compassion.

3.4.2. Anxiety

The main effect of time was significant, $F(1, 87) = 11.07, p = .001$, $\eta_p^2 = 0.113$. The main effect of group was not significant. The interaction effect between time and group was significant (see [Fig. 3](#)), $F(2, 87) = 10.43, p < .001, \eta_p^2 = 0.193$.

Simple effects analysis showed that anxiety significantly decreased after the intervention in Experimental Group 1, $F(1, 87) = 30.83, p < .001, \eta_p^2 = 0.262$. The Experimental Group 2 and Active Control Group showed no significant change. Additionally, post-intervention anxiety in Experimental Group 1 were significantly lower than those in Active Control Group ($p = .008$). There was no significant difference between

Table 2The NLP results of Total Word Count and EVs. ($df = 88$).

Variables	Negative event writing task		<i>t</i>	<i>p</i>	Writing intervention task		<i>t</i>	<i>p</i>
	Exp.	Act.			Exp.	Act.		
Total Word Count	198.81	180.23	0.76	0.446	210.91	164.47	1.28	0.203
Positive EVs	0.74	1.09	-1.94	0.055	0.99	0.51	2.05	0.042
Negative EVs	1.26	1.46	-0.99	0.326	0.97	0.03	5.70	<0.001

Notes: $p < 0.05$ defined statistical significance, same as below.**Table 3**

Demographic variables and descriptive statistics of the sample.

Variables	Exp. 1.	Exp. 2.	Act.	Statistics	<i>p</i>
Gender (n, %)				χ^2	
Male	10 (31.3)	10 (31.3)	12 (37.4)	0.388	0.824
Female	20 (34.5)	20 (34.5)	18 (31)		
Age (n, %)				χ^2	
18–24	16 (42.1)	10 (26.3)	12 (31.6)	8.292	0.405
25–29	2 (33.3)	3 (50)	1 (16.7)		
30–39	0 (0)	0 (0)	2 (100)		
40–49	10 (30.3)	13 (39.4)	10 (30.3)		
>50	2 (18.2)	4 (36.4)	5 (45.5)		
Baseline (<i>M</i> , <i>SD</i>)				<i>F</i>	
Self-Compassion	3.63	3.57 (0.43)	3.59 (0.75)	0.09	0.918
Anxiety	37.50	37.83 (9.3)	38.47 (12.13)		
Positive Emotion	34.17	36.80	37.57	2.30	0.106
Negative Emotion	(6.75)	(6.37)	(6.17)		
Mid-Assessment (<i>M</i> , <i>SD</i>)				<i>F</i>	
Positive Emotion	32.50	33.53	36.03	1.79	0.173
Negative Emotion	(7.94)	(7.35)	(6.98)		
Post-Assessment (<i>M</i> , <i>SD</i>)				<i>F</i>	
Self-Compassion	3.93	3.72 (0.4)	3.60 (0.79)	2.37	0.099
Anxiety	(0.58)				
Positive Emotion	32.20	37.03	39.07	3.88	0.024
Negative Emotion	(6.89)	(10.25)	(11.66)		
Emotion	34.80	36.33 (7.2)	36.20 (6.11)	0.45	0.637
Emotion	(7.35)				
Emotion	18.00	18.50	20.30	0.79	0.459
Emotion	(6.13)	(7.11)	(8.91)		

Experimental Group 1 and 2, or between Experimental Group 2 and Active Control Group.

Results indicate that self-compassion writing intervention with ChatGPT significantly reduces participants' anxiety.

3.5. Secondary results: positive and negative affect

Repeated measures ANOVA was conducted on positive and negative affect, with time (pre-intervention, post-negative event writing, post-writing intervention) as the within-subjects factor and group (Experimental Group 1, 2, and Active Control Group) as the between-subjects factor.

For the positive affect, the main effect of time was significant, $F(2, 87) = 8.17$, $p < .001$, $\eta_p^2 = 0.086$. The main effect of group and the interaction effect between time and group were not significant. Simple effects analysis showed that the simple effect of time was significant for Experimental Group 1, $F(2, 87) = 4.46$, $p = .014$, $\eta_p^2 = 0.094$, but not for Experimental Group 2 or Active Control Group. Positive affect significantly decreased after the negative event writing task for Experimental Group 1 ($p = .004$) and 2 ($p = .034$). Positive affect significantly increased after the self-compassion writing intervention for Experimental Group 1 ($p = .035$) and 2 ($p = .011$). There were no significant differences for other measures.

For the negative affect, the main effect of time was significant, $F(2,$

87) = 23.29, $p < .001$, $\eta_p^2 = 0.211$. The main effect of group and the interaction effect between time and group were not significant. Simple effects analysis showed that the simple effect of time was significant for Experimental Group 1, $F(2, 87) = 11.37$, $p < .001$, $\eta_p^2 = 0.209$, and for Experimental Group 2, $F(2, 87) = 6.67$, $p = .002$, $\eta_p^2 = 0.134$, but not for Active Control Group. Negative affect significantly increased after the negative event writing task for Experimental Group 1 ($p = .023$) and 2 ($p = .007$). Negative affect significantly decreased after the self-compassion writing intervention for Experimental Group 1 ($p < .001$) and 2 ($p < .001$), and also significantly decreased in the Control Group after the neutral writing task ($p = .024$). There were no significant differences for other measures.

Results indicate that regardless of ChatGPT involvement, negative event writing significantly increases negative affect and reduces positive affect, and self-compassion writing intervention significantly decreases negative affect and increases positive affect. ChatGPT's involvement did not significantly affect changes in positive and negative affect. Additionally, neutral writing also significantly decreases negative affect.

4. Discussion

Previous studies have demonstrated the effectiveness of self-compassion writing interventions in alleviating psychological issues such as depression and anxiety (Harwood and Kocovski, 2017; Wong and Mak, 2016). However, to the best of the author's knowledge, only one study have demonstrated the effectiveness of single online self-compassion writing interventions in reducing anxiety (Guan et al., 2021), and there is still no research validating ChatGPT's effectiveness in these interventions. We further confirmed the effectiveness and explored a new intervention method that did not require the involvement of professional therapists, which broadens the research outcomes and provides direction for future studies.

In this study, we observed significant increases in self-compassion and reductions in anxiety following the intervention. However, several discrepancies were noted when compared to prior self-compassion writing interventions. First, Experimental Group 2 did not exhibit significant reductions in anxiety, which contrasts with findings from previous studies. While several studies have demonstrated the efficacy of self-compassion writing in reducing anxiety (McEwan et al., 2018; Stevenson et al., 2019; Northover et al., 2021), these interventions typically involved continuous (2–8 weeks) and guided formats. The attenuated effect observed in the current study may be attributed to the single-session, online, self-help nature of the intervention, as suggested by Egan et al. (2022). This format may have led to a weaker impact on anxiety reduction in Experimental Group 2. Second, although no significant interaction effects were found for changes in emotional states, post-hoc analyses revealed that the experimental group experienced significant increases in positive affect (Cohen's $d = 0.41$) and decreases in negative affect (Cohen's $d = 0.73$), indicating a trend toward improved emotional states. However, it is challenging for single interventions to produce substantial changes in self-reported emotional indicators in the short term. This observation is consistent with Guan et al. (2021), who also reported no significant reduction in negative emotions following a single self-compassion writing intervention. Moreover, participants demonstrated notable changes in their emotional vocabulary (EVs) following the self-compassion writing task, suggesting

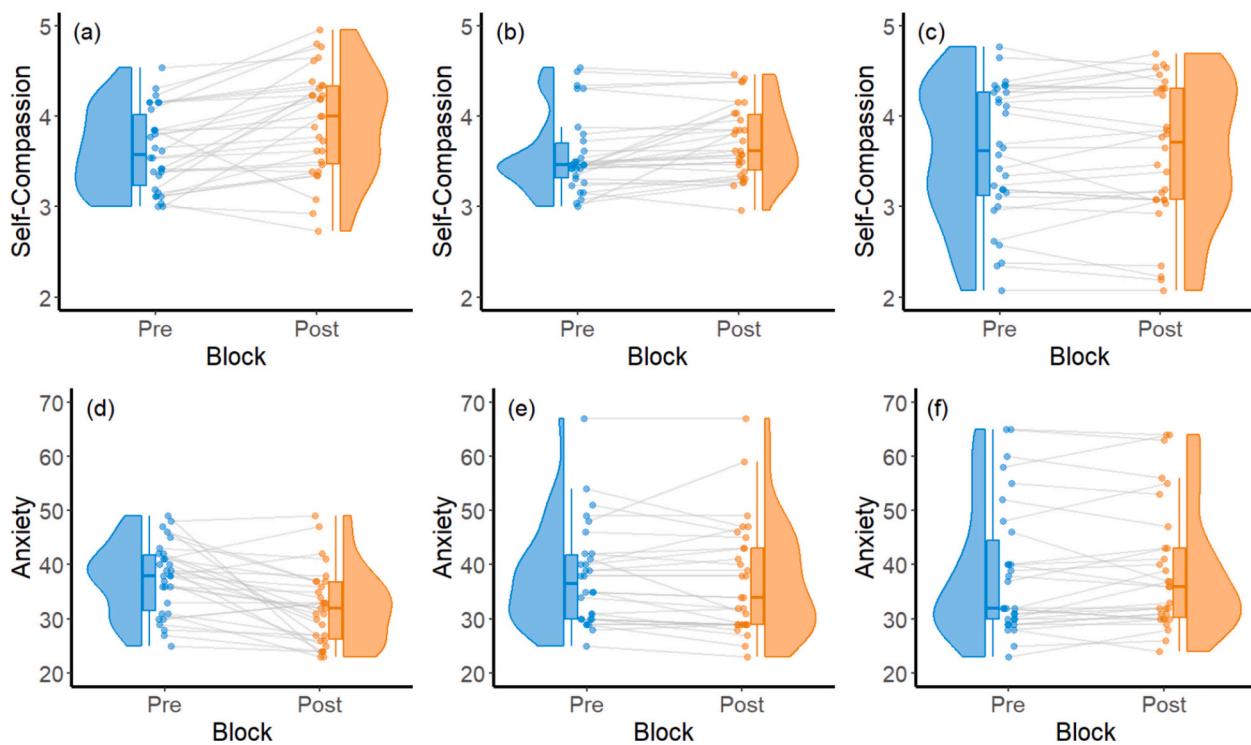


Fig. 3. Changes in Self-Compassion and Anxiety Before and After the Intervention.

(a) ~ (c) and (d) ~ (f) respectively show the changes in self-compassion and anxiety for Experimental Group 1, 2, and Active Control Group.

increased emotional awareness and regulation (Inwood and Ferrari, 2018), which aligns with the mindfulness aspects inherent to self-compassion. However, given that this was a brief, low-intensity online intervention, it is plausible that participants did not fully engage their emotional regulation skills. As a result, emotional awareness and regulation may not have been fully activated, which could explain the lack of significant changes in emotional indicators. Furthermore, we also acknowledge that the marginal significance of positive EVs in the negative event writing task for both the experimental and active control groups ($p = .055$) may seem counterintuitive. However, we believe this is in line with the inherent variability and instability of positive EVs (Vine et al., 2020). While this result does not alter our primary conclusions, it highlights the complex nature of measuring positive EVs and the need for further research to better understand their response to psychological interventions.

Although the single-session intervention showed some limitations in producing significant changes in anxiety and emotional states, it nevertheless demonstrated potential for enhancing self-compassion and emotional regulation. These findings suggest that, as a self-help tool, this type of intervention could be effective in fostering emotional awareness and self-compassion, particularly when implemented repeatedly over time. If used regularly, such interventions may have considerable potential to support individuals in managing their emotional well-being, offering a low-cost and accessible means of psychological support. Future research should explore the effects of repeated sessions to better understand the long-term benefits and efficacy of such self-help interventions.

Traditionally, research in clinical psychology or psychiatry has focused on the mechanisms, interventions, and outcomes for patients with anxiety disorders, often overlooking the much larger population of healthy individuals (Szuhany and Simon, 2022). Recently, the COVID-19 and economic downturn have intensified social competition and survival pressures, highlighting the importance of mental health. There is now increased attention on daily mental health care and prevention of mental illnesses for the general population (Salari et al., 2020). Against

this backdrop, we shift the focus from clinical populations to the broader goal of maintaining public mental health. We aim to design convenient, effective, and low-cost interventions to help people manage and regulate their emotions, prevent the onset of mental disorders in healthy individuals, reduce the likelihood of disease progression, and enhance life satisfaction and social well-being.

With the maturity of internet technology, convenient and fast online self-help interventions can effectively alleviate anxiety (Charbonnier et al., 2022). As a single-session intervention for the general population, self-compassion writing intervention demonstrates unique advantages in quickly and cost-effectively reducing public anxiety for preventive purposes. We conducted this study online, confirming its effectiveness in alleviating anxiety. The use of NLP validated the effectiveness, providing a reference for its broader application. Online self-help interventions are simple and effective, saving human and material resources while allowing more people to access mental health support anytime and anywhere (van Straten et al., 2008). By adapting self-compassion writing interventions for online self-help, we can reduce implementation costs, increase accessibility, and promote the efficient allocation of mental health resources (Kazdin, 2019).

Since 2023, ChatGPT has provided a new direction for clinical psychological interventions. In the field of psychotherapy, ChatGPT shows promise as a reliable assistant, offering additional options for individuals in need of psychological support. For example, trained to simulate professional psychologists, ChatGPT can accurately recognize emotions and provide appropriate responses (Elyoseph et al., 2023). Furthermore, research has shown that ChatGPT-generated responses indicate higher levels of authenticity, professionalism, and practicality than those produced by humans (Lopes et al., 2024), which suggest that ChatGPT-generated responses can be considered a reliable alternative to human-generated responses in certain applications.

ChatGPT has multiple potential applications in the field of mental health (Zhang and Zhao, 2024). One major application is as a therapeutic assistant to support the work of therapists. For example, ChatGPT can act as a therapist's assistant, communicating with patients,

collecting information, providing treatment suggestions, and recording therapy summaries and progress (Tal et al., 2023). It can also function as a companion support assistant, helping non-professionals offer more professional advice through collaborative human-computer interaction (Sharma et al., 2023). Another significant application is directly intervening with patients. We adopted this approach, using ChatGPT for the first time in self-compassion writing interventions. Under three elements of self-compassion, we focused on training ChatGPT to provide highly empathetic responses and aimed to use this empathy technique to enhance self-compassion and reduce anxiety, which may compensate for AI's lack of empathy and increase acceptance of AI applications in psychotherapy (Aktan et al., 2022). As an intervention mentor, ChatGPT provided timely feedback, resulting in better outcomes compared to the group without ChatGPT involvement. This further supports the findings of Carlbring et al. (2023), demonstrating the effectiveness of ChatGPT in clinical psychotherapy and validating ChatGPT's ability to empathise within a self-compassion framework, which providing empirical data for this emerging field.

Combining the results of this study and previous work (e.g. Jain et al., 2024; Lopes et al., 2024), AI has the potential to revolutionize mental health services by enhancing accessibility, scalability, and efficiency. AI-powered tools like ChatGPT can deliver real-time support using cognitive-behavioral techniques, psychoeducation, and emotional assistance, particularly in underserved areas or where stigma around seeking mental health care is prevalent. These tools provide immediate, non-judgmental responses that simulate the support of human therapists, enabling users to feel understood and supported. Additionally, AI can personalize interventions and track user progress, aligning with evidence-based psychological practices. However, it is important to view AI as a complement to human therapists, as it lacks the nuanced understanding required to address more severe or complex cases.

Additionally, participants in Experimental Group 1 were unaware that the feedback was generated by ChatGPT in this study, which may raise ethical considerations and implications for the results. Jain et al. (2024) found that when participants were informed about whether mental health intervention content was human- or AI-generated, human-generated content scored significantly higher on authenticity (Cohen's $d = 0.45$), while AI content showed no significant change. Meanwhile, trust in AI correlated significantly with its practicality rating ($r = 0.25$), reflecting a human preference for emotional intelligence and highlighting the role of trust in intervention effectiveness. At the same time, ethical concerns regarding AI in mental health support include risks of data privacy breaches, lack of transparency in data usage, potential over-reliance on AI, and its limitations in fully understanding complex emotions or resolving ethical dilemmas, which can lead to biased or inappropriate responses (Jain et al., 2024; Lopes et al., 2024). However, in this study, ChatGPT did not access user-identifiable data, and no sensitive information was collected, thereby minimizing the privacy-related risks. Future research should investigate how such disclosure might affect participants' perceptions of authenticity, trust, and overall intervention effectiveness, while also addressing broader ethical considerations in AI-driven mental health interventions.

There are still several limitations in our study. Our findings were based on single brief intervention sessions, and the long-term effectiveness remains unclear, requiring further follow-up studies for confirmation. Due to practical constraints, the experiment did not fully follow an RCT design, and PSM was used to minimize confounding effects from different experimental timings. Future research should increase sample size, improve RCT execution, consider cross-cultural consistency to enhance validity, and integrate cross-comparisons between ChatGPT's effectiveness and interventions guided by human therapists in different social, economic and cultural contexts. This will better minimize ethical risks, provide insight into the effectiveness and feasibility of AI using in mental health interventions, and evaluate user experience and limitations in various scenarios, thus optimizing collaboration and clarifying human and AI's respective roles and

boundaries. Additionally, we did not investigate the intervention's underlying mechanisms, which needs further investigation in future research.

Finally, with the continuous development of artificial intelligence, the rapid iteration and updates of ChatGPT may bring new ideas and methods to related intervention models. As an emerging research direction, exploring the application of ChatGPT in psychotherapy requires collaborative efforts from researchers in related fields. It is crucial to further explore and refine ChatGPT's uses in various online interventions and to develop diverse applications. Additionally, employing advanced training methods, such as Reinforcement Learning from Human Feedback (RLHF) and custom datasets, could enhance ChatGPT's capabilities in psychotherapy. We look forward to AI-driven online intervention programs and AI therapists better serving mental health needs and improving public well-being in the near future.

5. Conclusion

Online self-compassion writing intervention significantly increases self-compassion and positive affect, significantly reduces anxiety and negative affect; compared to self-compassion writing alone, participants with the feedback of ChatGPT achieved better intervention outcomes, with more pronounced increases in self-compassion and reductions in anxiety.

Funding

This research did not receive any specific grant from funding agencies.

CRediT authorship contribution statement

Tomoko Kishimoto: Conceptualization, Methodology, Validation, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Ximing Hao:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration. **Trimble Chang:** Conceptualization, Methodology, Software, Visualization. **Zhiye Luo:** Conceptualization, Methodology, Software, Formal analysis, Data curation, Visualization.

Declaration of competing interest

The authors do not have a conflict of interest (financial, academic, or otherwise) that would impede publication of this manuscript. In addition, this manuscript is original and has not been submitted (nor is it under consideration) for publication elsewhere.

Data availability

Data will be made available on request.

References

- Aktan, M.E., Turhan, Z., Dolu, I., 2022. Attitudes and perspectives towards the preferences for artificial intelligence in psychotherapy. *Comput. Hum. Behav.* 133, 107273. <https://doi.org/10.1016/j.chb.2022.107273>.
- Aydogdu, B.E., Dirik, G., 2022. Self-compassion focused writing practices: a systematic review. *Curr. Approach. Psychiatr.* 14 (4), 526–544. <https://doi.org/10.18863/pgy.1077952>.
- Baikie, K.A., Wilhelm, K., 2005. Emotional and physical health benefits of expressive writing. *Adv. Psychiatr. Treat.* 11 (5), 338–346. <https://doi.org/10.1192/apt.11.5.338>.
- Bandelow, B., Michaelis, S., 2015. Epidemiology of anxiety disorders in the 21st century. *Dialogues Clin. Neurosci.* 17 (3), 327–335. <https://doi.org/10.31887/DCNS.2015.17.3/bbandelow>.
- Brown, T.A., Campbell, L.A., Lehman, C.L., Grisham, J.R., Mancill, R.B., 2001. Current and lifetime comorbidity of the DSM-IV anxiety and mood disorders in a large

- clinical sample. *J. Abnorm. Psychol.* 110 (4), 585. <https://doi.org/10.1037/0021-843X.110.4.585>.
- Carlbring, P., Hadjistavropoulos, H., Kleiboer, A., Andersson, G., 2023. A new era in internet interventions: the advent of chat-GPT and AI-assisted therapist guidance. *Internet Interv.* 32. <https://doi.org/10.1016/j.invent.2023.100621>.
- Charbonnier, E., Trémolière, B., Baussard, L., Goncalves, A., Lespiau, F., Philippe, A.G., Le Vigouroux, S., 2022. Effects of an online self-help intervention on university students' mental health during COVID-19: a non-randomized controlled pilot study. *Comput. Human Behav. Rep.* 5, 100175. <https://doi.org/10.1016/j.chbr.2022.100175>.
- Chen, J., Yan, L., Zhou, L., 2011. Reliability and validity of Chinese version of self-compassion scale. *Chin. J. Clin. Psych.* 19 (6), 734–736. <https://doi.org/10.16128/j.cnki.1005-3611.2011.06.006>.
- Cox, R.C., Olatunji, B.O., 2016. A systematic review of sleep disturbance in anxiety and related disorders. *J. Anxiety Disord.* 37, 104–129. <https://doi.org/10.1016/j.janxdis.2015.12.001>.
- Craske, M.G., Rauch, S.L., Ursano, R., Prenoveau, J., Pine, D.S., Zinbarg, R.E., 2011. What is an anxiety disorder? *Focus* 9 (3), 369–388. <https://doi.org/10.1176/foc.9.3.foc369>.
- Dong, Z., Dong, Q., Hao, C., 2010. HowNet and its computation of meaning. In: *The proceedings of the 23rd international conference on computational linguistics: demonstrations* (pp. 53–56). <https://aclanthology.org/C10-3014.pdf>.
- Egan, S.J., Rees, C.S., Delalande, J., Greene, D., Fitzallen, G., Brown, S., Finlay-Jones, A., 2022. A review of self-compassion as an active ingredient in the prevention and treatment of anxiety and depression in young people. *Adm. Policy Ment. Health Ment. Health Serv. Res.* 49, 385–403. <https://doi.org/10.1007/s10488-021-01170-2>.
- Elyoseph, Z., Hadar-Shoval, D., Asraf, K., Lvovsky, M., 2023. ChatGPT outperforms humans in emotional awareness evaluations. *Front. Psychol.* 14, 1199058. <https://doi.org/10.3389/fpsyg.2023.1199058>.
- Gao, R., Hao, B.B., Li, H., Gao, Y.S., Zhu, T.S., 2013, October. Developing simplified Chinese psychological linguistic analysis dictionary for microblog. In: Paper Presented at the Meeting of 2013 International Conference on Brain & Health Informatics (BHI'13). Maebashi, Japan. https://doi.org/10.1007/978-3-319-02753-1_36.
- Guan, F., Wu, Y., Ren, W., Zhang, P., Jing, B., Xu, Z., He, J.B., 2021. Self-compassion and the mitigation of negative affect in the era of social distancing. *Mindfulness* 12, 2184–2195. <https://doi.org/10.1007/s12671-021-01674-w>.
- Harwood, E.M., Kocovski, N.L., 2017. Self-compassion induction reduces anticipatory anxiety among socially anxious students. *Mindfulness* 8, 1544–1551. <https://doi.org/10.1007/s12671-017-0721-2>.
- Huang, L., Yang, T., Ji, Z., 2003. Applicability of the positive and negative affect scale in Chinese. *Chin. Ment. Health J.* 1, 54–56. <https://doi.org/10.3321/j.issn:1000-6729.2003.01.018>.
- Inwood, E., Ferrari, M., 2018. Mechanisms of change in the relationship between self-compassion, emotion regulation, and mental health: a systematic review. *Appl. Psychol. Health Well Being* 10 (2), 215–235. <https://doi.org/10.1111/aphw.12127>.
- Jain, G., Pareek, S., Carlbring, P., 2024. Revealing the source: how awareness alters perceptions of AI and human-generated mental health responses. *Internet Interv.* 36, 100745. <https://doi.org/10.1016/j.invent.2024.100745>.
- Januzzi, J.L., Stern, T.A., Pasternak, R.C., DeSanctis, R.W., 2000. The influence of anxiety and depression on outcomes of patients with coronary artery disease. *Arch. Intern. Med.* 160 (13), 1913–1921. <https://doi.org/10.1001/archinte.160.13.1913>.
- Kazdin, A.E., 2019. Annual research review: expanding mental health services through novel models of intervention delivery. *J. Child Psychol. Psychiatry* 60 (4), 455–472. <https://doi.org/10.1111/jcpp.12937>.
- Ku, L.W., Liang, Y.T., Chen, H.H., 2006. Opinion extraction, summarization and tracking in news and blog corpora. In: *The proceedings of AAAI spring symposium: computational approaches to analyzing weblogs* (pp. 100–107). <https://cdn.aaai.org/Symposia/Spring/2006/SS-06-03/SS06-03-020.pdf>.
- Leary, M.R., Tate, E.B., Adams, C.E., Allen, A.B., Hancock, J., 2007. Self-compassion and reactions to unpleasant self-relevant events: the implications of treating oneself kindly. *J. Pers. Soc. Psychol.* 92 (5), 887–904. <https://doi.org/10.1037/0022-3514.92.5.887>.
- Leonard, K., Abramovitch, A., 2019. Cognitive functions in young adults with generalized anxiety disorder. *Eur. Psychiatry* 56 (1), 1–7. <https://doi.org/10.1016/j.eurpsy.2018.10.008>.
- Li, J., 2011. Chinese Polarity Lexicon v1.0. Retrieved May 17, 2024, from. <http://nlp.csai.tsinghua.edu.cn/news/112/>.
- Li, W., Qian, M., 1995. Revision of the state trait anxiety inventory (STAI) norm for Chinese university students. *Acta Sci. Nat. Univ. Pekin.* 1, 108–112. <https://doi.org/10.13209/j.0479-8023.1995.014>.
- Lopes, E., Jain, G., Carlbring, P., Pareek, S., 2024. Talking mental health: a battle of wits between humans and AI. *J. Technol. Behav. Sci.* 9, 628–638. <https://doi.org/10.1007/s41347-023-00359-6>.
- Lutz, W., Schiefele, A.K., Wucherpfennig, F., Rubel, J., Stulz, N., 2016. Clinical effectiveness of cognitive behavioral therapy for depression in routine care: a propensity score based comparison between randomized controlled trials and clinical practice. *J. Affect. Disord.* 189, 150–158. <https://doi.org/10.1016/j.jad.2015.08.072>.
- Mazza, M.G., De Lorenzo, R., Conte, C., Poletti, S., Vai, B., Bollettini, I., COVID-19 BioB Outpatient Clinic Study Group, 2020. Anxiety and depression in COVID-19 survivors: role of inflammatory and clinical predictors. *Brain Behav. Immun.* 89, 594–600. <https://doi.org/10.1016/j.bbi.2020.07.037>.
- McCarthy, J.M., Trougakos, J.P., Cheng, B.H., 2016. Are anxious workers less productive workers? It depends on the quality of social exchange. *J. Appl. Psychol.* 101 (2), 279. <https://doi.org/10.1037/apl0000044>.
- McEwan, K., Elander, J., Gilbert, P., 2018. Evaluation of a web-based self-compassion intervention to reduce student assessment anxiety. *Interdisc. Ed. Psychol.* 2 (1), 6. <https://derby.openrepository.com/handle/10545/622811>.
- Mei, Y., Liu, J., Liu, H., Fu, Y., Luo, X., Lei, Y., 2023. Self-compassion writing facilitates fear extinction. *Acta Psychol. Sin.* 8, 1317–1329. <https://doi.org/10.3724/SP.J.1041.2023.01317>.
- Montag, C., Becker, B., Gan, C., 2018. The multipurpose application WeChat: a review on recent research. *Front. Psychol.* 9, 408296. <https://doi.org/10.3389/fpsyg.2018.02247>.
- Nazarova, D., 2023. Application of artificial intelligence in mental healthcare: Generative pre-trained transformer 3 (GPT-3) and cognitive distortions. In: *Proceedings of the Future Technologies Conference*. Springer Nature Switzerland, Cham, pp. 204–219. https://doi.org/10.1007/978-3-031-47454-5_16, October.
- Necho, M., Tsehay, M., Birke, M., Biset, G., Tadesse, E., 2021. Prevalence of anxiety, depression, and psychological distress among the general population during the COVID-19 pandemic: a systematic review and meta-analysis. *Int. J. Soc. Psychiatry* 67 (7), 892–906. <https://doi.org/10.1177/00207640211003121>.
- Neff, K.D., 2003. Self-compassion: an alternative conceptualization of a healthy attitude toward oneself. *Self Identity* 2 (2), 85–101. <https://doi.org/10.1080/15298860309032>.
- Northover, C., Deacon, J., King, J., Irons, C., 2021. Developing self-compassion online: assessing the efficacy and feasibility of a brief online intervention. *OBM Integr. Complement. Med.* 6 (4), 1–29. <https://doi.org/10.21926/obm.icm.2104056>.
- Odou, N., Brinker, J., 2015. Self-compassion, a better alternative to rumination than distraction as a response to negative mood. *J. Posit. Psychol.* 10 (5), 447–457. <https://doi.org/10.1080/17439760.2014.967800>.
- Powell, T., Enright, S., 2015. Anxiety and stress management. Routledge. <https://doi.org/10.4324/9781315683867>.
- Salari, N., Hosseiniyan-Far, A., Jalali, R., Vaisi-Raygani, A., Rasoulopoor, S., Mohammadi, M., Khaledi-Paveh, B., 2020. Prevalence of stress, anxiety, depression among the general population during the COVID-19 pandemic: a systematic review and meta-analysis. *Glob. Health* 16 (1), 1–11. <https://doi.org/10.1186/s12992-020-00589-w>.
- Schleider, J.L., Weisz, J.R., 2017. Little treatments, promising effects? Meta-analysis of single-session interventions for youth psychiatric problems. *J. Am. Acad. Child Adolesc. Psychiatry* 56 (2), 107–115. <https://doi.org/10.1016/j.jaac.2016.11.007>.
- Sharma, A., Lin, I.W., Miner, A.S., Atkins, D.C., Althoff, T., 2023. Human-AI collaboration enables more empathic conversations in text-based peer-to-peer mental health support. *Nat. Machine Intel.* 5 (1), 46–57. <https://doi.org/10.1038/s42256-022-00593-2>.
- Spielberger, C.D., Gorsuch, R.I., Lushene, R.E., 1970. STAI Manual for the State-Trait Anxiety Inventory. Consulting Psychologist Press, Palo Alto, CA. <https://doi.org/10.30849/rip.v5i3%20&%204.620>.
- Stevenson, J., Mattiske, J.K., Nixon, R.D., 2019. The effect of a brief online self-compassion versus cognitive restructuring intervention on trait social anxiety. *Behav. Res. Ther.* 123, 103492. <https://doi.org/10.1016/j.br.2019.103492>.
- Szuhany, K.L., Simon, N.M., 2022. Anxiety disorders: a review. *Jama* 328 (24), 2431–2445. <https://doi.org/10.1001/jama.2022.22744>.
- Tal, A., Elyoseph, Z., Haber, Y., Angert, T., Gur, T., Simon, T., Asman, O., 2023. The artificial third: utilizing ChatGPT in mental health. *Am. J. Bioeth.* 23 (10), 74–77. <https://doi.org/10.1080/15265161.2023.2250297>.
- van Straten, A., Cuijpers, P., Smits, N., 2008. Effectiveness of a web-based self-help intervention for symptoms of depression, anxiety, and stress: randomized controlled trial. *J. Med. Internet Res.* 10 (1), e954. <https://doi.org/10.2196/jmir.954>.
- Vine, V., Boyd, R.L., Pennebaker, J.W., 2020. Natural emotion vocabularies as windows on distress and well-being. *Nat. Commun.* 11 (1), 4525. <https://doi.org/10.1038/s41467-020-18349-0>.
- Watson, D., Clark, L.A., Tellegen, A., 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. *J. Pers. Soc. Psychol.* 54 (6), 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>.
- Wong, C.C., Mak, W.W., 2016. Writing can heal: effects of self-compassion writing among Hong Kong Chinese college students. *Asian Am. J. Psychol.* 7 (1), 74–82. <https://doi.org/10.1037/aap0000041>.
- Zhang, D., Zhao, X., 2024. Understanding adoption intention of virtual medical consultation systems: perceptions of ChatGPT and satisfaction with doctors. *Comput. Hum. Behav.* 159, 108359. <https://doi.org/10.1016/j.chb.2024.108359>.