



ARTICLE

The Intervention Effects of Self-Perspective Imaginative Empathy on Mental Illness Stigma in Chinese Undergraduates

Siliang Yang^{1,2,*} and Gan Jin³

¹School of Educational Science, Anhui Normal University, Wuhu, China

²Tongcheng Teachers College, Tongcheng, China

³College of Education, Washington State University, Pullman, WA, USA

*Corresponding Author: Siliang Yang. Email: ahnuyang@tctc.edu.cn

Received: 07 January 2026; Accepted: 10 April 2026; Published: 30 July 2026

ABSTRACT: Background: Reducing mental illness stigma is critical for mental health literacy enhancement in China, where stigmatization remains pronounced among undergraduates. Imaginative empathy is a promising anti-stigma intervention, but its specific effects and the potential moderating role of gender remain under-explored. **Methods:** Three studies were conducted with 18–22-year-old Chinese undergraduates. Study 1 ($n = 223$, survey design) explored gender's moderation on empathy-stigma links. Study 2 ($n = 217$, $2 \times 2 \times 3$ mixed design) tested the intervention effect of self-perspective imaginative empathy. Study 3 ($n = 71$, $2 \times 2 \times 2$ mixed design) verified gender moderation. Stigma was measured via stereotypic explanatory bias (SEB) with internal attribution tendency as the core indicator. **Results:** Gender moderated the empathy-stigma relationship in positive situations (Study 1: $\beta = 0.82$, $p < 0.001$; Study 3: $F = 4.77$, $p = 0.032$), with males showing increased internal attribution as empathy rise. Self-perspective imaginative empathy significantly increased internal attribution in positive situations ($F = 8.70$, $p < 0.01$) and decreased it in negative situations ($F = 109.63$, $p < 0.001$; Study 2). **Conclusion:** Self-perspective imaginative empathy elicits attributional shifts consistent with reduced mental illness stigma among Chinese undergraduates. It induces a significant increase in internal attribution tendency among males in positive situations, providing implications for developing culturally tailored anti-stigma interventions targeting implicit cognitive biases.

KEYWORDS: Imaginative empathy; stigma; attitudinal interventions; gender moderation

1 Introduction

Mental health literacy refers to the knowledge, attitudes, and behaviors individuals develop to promote mental health and cope with mental illnesses, encompassing an understanding of positive mental health maintenance, mental disorders recognition and treatment, stigma reduction, and help-seeking ability enhancement [1]. Currently, Chinese college students face a dual challenge of high mental illness prevalence and low medical consultation rates [2]. A meta-analysis found the prevalence of internalized mental health problems among this group to be 17.7% [3], yet few students actively seek counseling, and most hold negative attitudes toward help-seeking [4]. An epidemiological survey (2001–2005) of 113 million adults showed only 8% of those with mental disorders sought professional help, and just 5% were seen by mental health professionals [5]. A primary driver of this phenomenon is the reluctance to seek help due to mental illness stigma, making stigma reduction a crucial priority for addressing college students' mental health disparities [6].

1.1 Measurement and Intervention of Mental Illness Stigma

Traditional mental illness stigma measures suffer from cumbersome administration and social desirability bias [7]. To address this, Hastie's explanatory bias [8] was refined into the stereotypic explanatory bias (SEB)—an indicator of implicit attitude measurement reflecting how social group stereotypes shape information processing [9]. Rooted in attribution theory, SEB posits that individuals attribute stereotype-consistent behaviors to internal factors (personal traits) and stereotype-inconsistent behavior to external factors (environment) [10]. SEB is contextualized, effectively reflects implicit attitudes, and has strong predictive validity [9,11], making it an ideal tool for measuring mental illness stigma.

The essence of mental illness stigma is the social assignment of negative labels to individuals with mental illness, with their adverse circumstances attributed to stable, controllable personal deficits [12]. When processing behavioral of this group, individuals first infer causal factors and responsibility, which shapes subsequent responses [13]. Internal attribution tendency—the propensity to attribute behaviors to personal traits, abilities, or choices—serves as the core SEB indicator for measuring stigma. In positive situations: Higher internal attribution (PI-PE, positive internal minus external attribution) means recognizing the internal traits (e.g., ability, motivation) behind positive behaviors of individuals with mental illness, breaking the negative stereotype of “incapacity for positive action” and reflecting reduced stigma [14]. In negative situations: Higher internal attribution (NI-NE, negative internal minus external attribution) means blaming negative behaviors on personal deficits, reinforcing the stereotype of “personal problems causing adverse outcomes” and reflecting increased stigma [14]. Per the attribution-empathy model [15], empathic responses are reduced when negative behaviors are internally attributed or positive behaviors are external attributed; conversely, external attribution for negative behaviors or internal attribution for positive behaviors boosts empathy and alters emotional/behavioral tendencies toward individuals with mental illness. This makes internal attribution tendency a valid predictor of discriminatory emotional and behavioral tendencies [16], aligning with SEB measurement logic [9].

Mental illness stigma interventions are a core component of mental health literacy interventions, with targeted efforts implemented globally to improve public knowledge, attitudes, and behaviors [17]. A meta-analysis of Chinese mental health literacy interventions found only stigma reduction and help-seeking behavior enhancement to have a significantly positive association [18], with the proposition that “reducing stigmatizing attitudes promotes help-seeking behaviors” widely supported [19]. In Chinese culture, mental illness is often misperceived as a sign of weakness or lack of willpower [20], and collectivist cultural norms amplify stigmatization—even extending to the patient's family and social circles [21]. Thus, mental health literacy promotion in China currently centers on stigmatizing attitude interventions.

1.2 Imaginative Empathy

Empathy is a multifaceted psychological construct involving emotional resonance, cognitive understanding, and behavioral responses, defined as the abilities/traits to feel and understand others' emotions and exhibit benevolence and assistance [22]. Stotland pioneered guided imagery to evoke empathy in experiments [23], a method now termed imaginative empathy [24] which has been operationalized into two distinct forms with notable psychological and neuroscientific differences [23,24]. Self-perspective (imagining oneself in others' situations) activates emotion-related brain regions more significantly [25], fostering profound emotional empathy and insight into others' dilemmas; other-perspective (imagining others' feelings) is a cold perceptual/cognitive process that rarely evokes emotional resonance [24].

Empathy induction effectively improves attitudes towards stigmatized outgroups (e.g., the elderly, individuals with mental illness, people with AIDS, and homeless) [26]. While the intergroup contact

hypothesis posits that social contact reduces prejudice [27], non-contact interventions yield significantly stronger anti-stigma effects in China [18], as low-quality contact can even exacerbate stigma [28]. Imagination shares the same neural basis as perception and memory [29] and generates emotional processes analogous with real situations [30], making imaginative empathy a scientific and effective non-contact anti-stigma intervention.

The Self-Expansion Theory posits that individuals have an intrinsic motivation to incorporate external resources (including others) into the self to enhance their abilities, improve well-being, and maintain relationships [31]. When individuals experience psychological overlap with a group, they tend to view the group's members as "us" rather than "them" [32]. When adopting the perspective of individuals with mental illness (i.e., self-perspective imaginative empathy), individuals activate self-related cognitive representations to understand their experiences [33]. This enhances self-other overlap, which is rooted in shared neural circuits and integrated cognitions [34], and creates a conflict with existing stigmatizing stereotypes [35]. The conflict modulates cognitive resource allocation, weakening stereotypical accessibility and fostering more nuanced perceptions [36]. In contrast, other-perspective imaginative empathy (e.g., objectively analyzing the thoughts and feelings of the target) is more of a cold cognitive process [37], which hardly evokes emotional resonance with the predicament of people with mental illness. As a result, individuals lose the motivation to transcend their own positions and recognize the behaviors of people with mental illness. Cognitive analysis from the other-perspective may even strengthen individuals' implicit cognitive stigma by making them focus more on group differences [38].

Hypothesis 1: *Self-perspective imaginative empathy has a positive effect on mental illness stigma interventions.*

1.3 Moderating Effects of Gender

Gender differences in empathic responses are a key focus of empathy research [39], with a scholarly consensus that females have an inherent advantage in emotion recognition and exhibit higher levels than males [40]. These differences intensify from infancy to adolescence, driven by pubertal gonadal system activation. The secretion of oxytocin by the female body has been demonstrated to enhance individual empathic responses [41]. In contrast, the androgenic hormone testosterone in males has been shown to attenuate empathic responses [42]. Oxytocin more often facilitates males' social cognition and positive social interactions [43]. Moreover, it has been demonstrated to facilitate enhanced empathic ability, characterized by a reduction in egocentric tendencies, an improvement in perspective-taking skills, and an elevated capacity to experience and attend to the emotions of others [41]. These effects suggest that females exhibit higher levels of physiologically based trait empathy [40], maintain more positive perceptions and attitudes, and demonstrate reduced tendencies to stigmatize. A stigma intervention study among individuals taking antidepressants demonstrated that perspective-taking was more effective in improving stigmatizing attitudes in males than in females, possibly because female participants had higher levels of trait-based empathy and thus already exhibited high levels of perspective-taking with individuals on antidepressants prior to receiving the intervention [44].

Hypothesis 2: *Gender moderates the relationship between imaginative empathy and stigmatizing attitudes toward mental illness.*

2 Study 1: Moderating Effects of Gender in the Relationship between Empathy and Stigmatizing Attitudes

2.1 Methods

2.1.1 Participants and Design

Before introducing the participants of Study 1, a general overview of the sample characteristics across the three studies is presented in Table 1. All studies recruited Chinese undergraduate students aged 18–22 years, with consistent inclusion and exclusion criteria. It is critical to emphasize that the participants employed in the three studies are fully independent—no participant took part in more than one study. To ensure this independence, we implemented rigorous screening procedures: participants were selected from distinct universities and academic majors across the three studies, and all potential participants were queried during recruitment to confirm they had not participated in any similar studies on mental illness stigma or empathy prior to recruitment. The Academic Ethics Committee of Anhui Normal University, China (Approval No.: AHNU-ET2022002) approved the study. Consent was obtained from the students themselves before data collection, allowing all participants to withdraw at any time during the research process.

Table 1: Basic Characteristics of Participants Across Three Studies.

Variable	Study 1	Study 2	Study 3
Final sample size	223	217	71
Males, n	76	74	35
Females, n	147	143	36
Age (years), Mean $\pm$ SD	20.05 $\pm$ 0.99	20.04 $\pm$ 1.00	19.96 $\pm$ 0.99
Recruitment setting	Mental health public courses	Mental health public courses	Campus information platform
Major fields	Preschool Education, Art & Design, Applied Chemistry	Education, Civil Engineering, Music Performance	Education, Mathematics, Broadcasting

A random sample of 240 undergraduate students was selected from the college public course on mental health for the study, covering multiple academic fields including Preschool Education, Advertising Art and Design, and Applied Chemistry. The inclusion criteria were set as: aged 18 years or above, self-reported no history of severe psychological or mental disorders, and voluntary participation with signed informed consent forms. The exclusion criteria included: currently receiving psychological counseling or therapy, concurrently participating in other research projects, and inability to understand the task requirements. 7 cases of invalid data were excluded due to missing values, which resulted in a final sample of 223 participants (76 males, 147 females) aged between 18 and 22 (Mean = 20.05, SD = 0.99), among whom 108 were from urban areas and 115 from rural areas. The distribution characteristics of gender, age and household registration statuses were consistent with the natural distribution of undergraduate students in the sampling frame. After completing the registration of basic information, including gender and age, the participants completed the empathy measure and the stigmatizing attitude test.

2.1.2 Research Instruments

Empathy Measurement

The Social Empathy Competence Questionnaire [45], a scale developed based on Chinese cultural background and the characteristics of Chinese students with good reliability and validity, was utilized to assess the participants' capacity to accurately perceive and evaluate the emotions, feelings, and cognitions of others in social contexts. The correlation coefficients between this questionnaire and the two dimensions

of the Basic Empathy Scale (cognitive empathy and affective empathy) were 0.41 and 0.48, respectively, indicating satisfactory criterion-related validity. In the present study, the scale achieved a Cronbach's α coefficient of 0.77 and a McDonald's ω coefficient of 0.77, showing acceptable internal consistency. The questionnaire consists of 18 items, with 1 point awarded for each accurate response. The total score was converted into a standardized score ranging from 1 to 5, where scores of 3 or below indicate a low or average level of social empathy ability, and scores of 4 or above indicate a satisfactory level of social empathy ability.

Stigma Attitude Test

The behavioral attribution task was designed according to the SEB research paradigm [9]. Firstly, conceptual terms related to individuals with mental illnesses were selected from previous studies, and 100 Chinese undergraduates who were not participants were asked to rate the representativeness of the conceptual terms on a 5-point scale. Eight items related to individuals with mental illnesses were identified, which included obsessive-compulsive disorder (OCD) patients, phobia patients, addiction disorder patients, and neurasthenia patients. The second stage of the study involved the design of neutral protagonist titles, such as "Bai, Yang, Song", etc. The third stage involved the construction of eight event outcomes, including four positive event outcome descriptions and four negative event outcome descriptions. Combined with the eight-item subjects and protagonist titles to form declarative sentences. Finally, the eight test tasks were mixed with neutral statements and randomly ordered for participants to complete in order. The tasks required participants to visualize the statements describing the situation and to write down possible reasons for the outcome.

This attribution task has been formally applied to the measurement of mental illness stigma in Chinese undergraduate samples in previous research, which verified its sound psychometric properties and recognized its ability to effectively assess implicit internal attribution tendencies associated with mental illness stigma. Specifically, it demonstrated a 4-week test-retest reliability of 0.86 and a Cronbach's α coefficient of 0.80 [46], indicating good temporal stability and internal consistency, and thus serving as a valid measurement tool in the present study.

Responses were screened to exclude non-causal explanations, then coded for internal/external attribution by two independent, double-blind (blind to gender/experimental group) psychological researchers. Multiple explanations for a single event were coded separately, and the number of internal/external attribution elements was counted as PI (positive internal), PE (positive external), NI (negative internal), NE (negative external). Inter-rater reliability was excellent ($ICC = 0.89$, $p < 0.001$); ambiguous/meaningless responses (2.1% invalid rate) were jointly judged and excluded. Internal attribution tendency was calculated as PI-PE (positive situations) and NI-NE (negative situations), with scores interpreted as outlined in Section 1.1.

2.1.3 Statistical Analysis

All data analyses were conducted using SPSS 27.0 (IBM Corp., Armonk, NY, USA). The statistical tests employed included an independent-samples t -test, hierarchical regression analysis, and simple slope analysis. All statistical tests were two-tailed, and the significance level was set at $\alpha = 0.05$ for all analyses. Missing data were handled by listwise deletion, and invalid responses (e.g., regular answering patterns, irrelevant causal explanations) were excluded from the final analysis. Simple slope analysis was performed to further verify the moderating effect of gender, and the results were visualized with simple slope plots.

2.2 Results

2.2.1 Gender Differences in Stigma Attitudes

Independent samples *t*-tests compared the internal attribution tendency of male and female participants in positive and negative situations, respectively (Table 2). In positive situations, female undergraduates exhibited a significantly higher internal attribution tendency than males ($t = -2.14$, $p = 0.03$, Cohen's $d = -0.30$), reflecting a small effect size. In negative situations, no significant gender difference was observed ($t = 0.70$, $p = 0.48$, Cohen's $d = 0.10$), indicating a negligible effect with no practical significance.

Table 2: Gender differences in stigma attitudes (Mean $\pm$ SD).

Situation	Male	Female	<i>t</i>	<i>p</i>	Cohen's <i>d</i>
Positive	0.45 $\pm$ 2.08	1.02 $\pm$ 1.80	-2.14*	0.03	-0.30
Negative	0.45 $\pm$ 1.43	0.29 $\pm$ 1.73	0.70	0.48	0.10

Note: * $p < 0.05$; Effect size benchmarks [47]: Small ($d \geq 0.20$), Moderate ($d \geq 0.50$), Large ($d \geq 0.80$).

2.2.2 Moderating Effects of Gender in the Relationship between Empathy and Stigma

The moderating effect of gender in the relationship between empathy and stigma was analyzed using hierarchical regression in positive and negative situations, respectively (Table 3).

Table 3: Hierarchical regression analyses.

Dependent Variables	Predictor	Level 1		Level 2	
		β	<i>t</i>	β	<i>t</i>
internal attribution tendency in positive situations	Intercept	0.83	6.47***	-0.17	-0.56
	Empathy	0.05	0.39	-0.21	-1.49
	Gender	-0.28	-2.15*	-1.64	-4.16***
	Empathy $\times$ Gender			0.82	3.65***
	R^2	0.02		0.08	
	ΔR^2	0.02		0.06	
	<i>F</i>	2.35		6.09***	
internal attribution tendency in negative situations	Intercept	0.34	3.11**	0.61	2.32*
	Empathy	-0.04	-0.32	0.04	0.29
	Gender	0.08	0.72	0.45	1.31
	Empathy $\times$ Gender			-0.23	-1.14
	R^2	0.00		0.01	
	ΔR^2	0.00		0.01	
	<i>F</i>	0.30		0.63	

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

Hierarchical regression analysis was conducted with positive situation internal attribution tendency as the dependent variable, (f^2 effect size benchmarks: small ≥ 0.02 , medium ≥ 0.15 , large ≥ 0.35) [47]. In Step 1, empathy and gender were entered as predictors: empathy had no significant predictive effect ($\beta = 0.05$, $p = 0.696$), while gender was a significant predictor ($\beta = -0.28$, $p = 0.032$). Step 2 added the empathy $\times$ gender interaction term, which significantly increased the model's explanatory power ($\Delta R^2 = 0.06$, $f^2 = 0.07$, F change = 13.30, $p < 0.001$; small-to-medium effect) and emerged as a significant predictor ($\beta = 0.82$, $p < 0.001$), indicating gender significantly moderated the empathy-internal attribution tendency relationship. To further validate the moderating effect, simple slope analyses were performed for different genders (male = 1, female = 0) and visualized in a simple slope plot (Fig. 1A). Results revealed that in positive

situations, male undergraduates' internal attribution tendency significantly increased with higher empathy (*simple slope* = 0.93, $p < 0.001$), whereas no significant change in internal attribution tendency was observed among female undergraduates as empathy increased (*simple slope* = 0.11, $p = 0.251$). Collectively, these findings demonstrate that gender significantly moderates the association between empathy and internal attribution tendency in Chinese undergraduates under positive situational contexts.

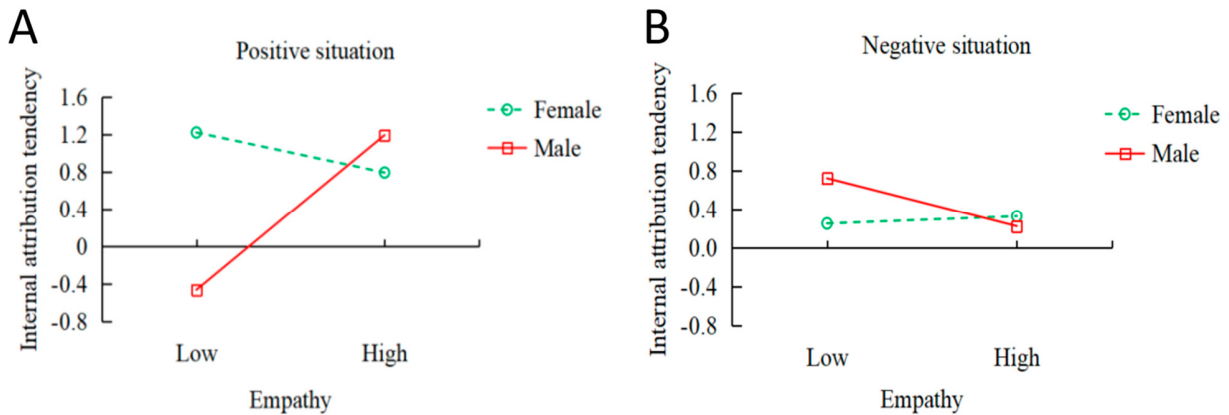


Figure 1: Simple slope plots of the interaction between Empathy and Gender on internal attribution tendency. (A) Positive situation; (B) Negative situation.

With participants' internal attribution tendency in negative situations as the dependent variable, a hierarchical regression analysis was performed. In Step 1, neither empathy ($\beta = -0.04$, $p = 0.749$) nor gender ($\beta = 0.08$, $p = 0.472$) significantly predicted internal attribution tendency. Step 2 included the empathy $\times$ gender interaction term, which did not lead to a meaningful increase in the model's explanatory power ($\Delta R^2 = 0.01$, $f^2 = 0.01$, F change = 1.29, $p = 0.258$). According to the f^2 benchmarks, this effect was below the threshold for a small effect and thus practically negligible, implying that the interaction term contributed almost no additional explanatory power to the model. The interaction failed to exert a significant predictive effect on internal attribution tendency ($\beta = -0.23$, $p = 0.258$). These results indicate that gender does not significantly moderate the relationship between empathy and internal attribution tendency among Chinese undergraduates in negative situations (Fig. 1B).

2.3 Discussion

Study 1 findings demonstrated that gender moderates the relationship between empathy and stigma attitudes in positive situations. Specifically, female undergraduates' internal attribution tendency remained unchanged when presented with positive behaviors of individuals with mental illness, with both low- and high-empathy females exhibiting high levels of internal attribution. Regarding gender differences in stigma (Table 2), female undergraduates showed significantly higher internal attribution tendency than their male counterparts. Puberty is characterized by increased oxytocin secretion in females, a hormone that enhances social interaction and emotional empathy [41], enabling them to perceive and attend to others' emotions and maintain more positive attitudes toward individuals with mental illness.

The hypothesis that gender moderates the association between empathy and stigma in negative situations was not supported. Neither male nor female undergraduates showed significant changes in internal attribution tendency when confronted with negative behaviors of individuals with mental illness, with both groups demonstrating low positive internal attribution levels. Female undergraduates' internal attribution tendency was marginally significantly different from 0 ($t = 2.01$, $p = 0.047$), while

male undergraduates' internal attribution tendency was significantly higher than 0 ($t = 2.73$, $p = 0.008$), indicating a more pronounced stigmatizing tendency among males. Males have higher testosterone levels, which impair the ability to perceive peers' thoughts and feelings in social interactions [42]. Consequently, physiologically based low trait empathy may constrain the reduction of internal attribution tendency (i.e., reduced stigma) among male undergraduates in negative situations.

From a sociocultural perspective, gender role socialization encourages females to be more emotionally expressive, nurturing, and interpersonally sensitive, which contributes to the development and expression of empathy. In contrast, males are socialized to emphasize independence and emotional restraint, which may reduce their attention to others' difficulties and weaken empathic responses. Meanwhile, gender-related social norms also shape different patterns of stigma perception. Females tend to show greater tolerance and understanding toward vulnerable groups, whereas males are more likely to hold relatively strong stereotypes under traditional gender role expectations. These sociocultural factors related to gender roles work together to make gender moderate the association between empathy and stigmatizing attitudes.

Overall, the findings indicate that gender significantly moderates the relationship between empathy and stigmatizing attitudes, particularly in the context of imagining positive behaviors of individuals with mental illness.

3 Study 2: Positive Effects of Imaginative Empathy on a Stigma Intervention

3.1 Methods

3.1.1 Participants and Design

The calculations conducted using G*Power 3.1 have indicated that for repeated measures ANOVA between participants, the expected number of participants required is a minimum of 54 (Effect size = 0.25, $\alpha = 0.05$, Power = 0.95). The study utilized a 2 (measurement time: pre-test vs. post-test, within-subjects) $\times$ 2 (situation: positive vs. negative, within-subjects) $\times$ 3 (group: self-perspective vs. other-perspective vs. control, between-subjects) experimental mixed design. A total of 223 undergraduate students were randomly selected via a random number table from the enrollment lists of the College Students' Mental Health Education course across three undergraduate universities, covering multiple academic fields, including Education, Civil Engineering, and Music Performance. The inclusion criteria were as follows: aged 18 years or above, self-reported no history of severe psychological or mental disorders, and willing to participate in the study with experimental intervention and sign the informed consent form. The exclusion criteria included currently receiving psychological counseling or therapy, concurrently participating in other research projects, and being unable to understand the task requirements. Participants were then allocated to the self-perspective group (25 males, 49 females), other-perspective group (25 males, 49 females), and control group (26 males, 49 females) using stratified random assignment with gender as the stratification variable. Six participants were excluded due to missing data, resulting in a final valid sample of 217 participants, consisting of 24 males and 49 females in the self-perspective group, 25 males and 47 females in the other-perspective group, and 25 males and 47 females in the control group. All participants were aged from 18 to 22 (Mean = 20.04, SD = 1.00). No significant intergroup differences were found among the three groups in terms of gender distribution ($\chi^2(2) = 0.07$, $p = 0.964$, Cramer's V = 0.02, 95% CI [0.01, 0.19]), age ($F(2, 214) = 1.08$, $p = 0.358$, partial $\eta^2 = 0.02$), pre-test scores of internal attribution tendency in positive situations ($F(2, 214) = 2.35$, $p = 0.073$, partial $\eta^2 = 0.03$), and pre-test scores of internal attribution tendency in negative situations ($F(2, 214) = 1.93$, $p = 0.126$, partial $\eta^2 = 0.03$) all showed no significant intergroup differences across the three groups.

3.1.2 Procedures

Following the registration of basic demographic information, including gender and age, the participants were administered the stigmatizing attitude pre-test (T1). The experimental procedure was conducted in groups, and the participants were invited to complete the empathy elicitation validity test and the stigmatizing attitude post-test (T2).

Stigma Test

Same as Study 1.

Experimental Manipulation Materials

Adopting the imagination paradigm [23,24], the experimental induction materials included textual and video components, both selected from authoritative, professional, and highly influential news platforms in China. The textual materials were selected from China Economic Weekly's Our Patients—A Report on China's Mental Patients (1422 Chinese characters) [48], and the video materials were from News 1+1 (Issue No. 20130506, titled *Mental Disorders are Cured, How to Cure Social Stigma?*) with a total duration of 25 min [49]. The content focused on the life experiences of individuals with mental illnesses and the pressure and discrimination they face from their families, friends, and society.

The experimental induction materials were reviewed and validated by three professors and associate professors in clinical and social psychology to ensure that the content accurately reflects the real experiences of individuals with mental illness without overstatement or stereotyping, and is capable of effectively evoking empathic responses. A 7-point Likert scale (1 = strongly disagree, 4 = neutral, 7 = strongly agree) was used to independently evaluate the text material and video material on the following four dimensions: Objectivity (the material is objective, without exaggeration or stigma, and truly reflects the real life of individuals with mental disorders), Emotional pleasantness (the emotional valence of the material), Emotional arousal (the intensity of emotion evoked by the material), Suitability (the content and duration are appropriate for the research topic and for Chinese undergraduates). The mean scores of the text material were 7.00 for objectivity, 1.67 for pleasantness, 5.67 for arousal, and 6.67 for suitability. The mean scores of the video material were 7.00 for objectivity, 1.33 for pleasantness, 5.33 for arousal, and 6.67 for suitability. The intraclass correlation coefficients (ICC) for the two materials were 0.98 and 0.99, respectively. The results indicated that both the text and video materials had good objectivity, emotional appropriateness, and applicability, and could be used as effective experimental materials for the present study.

The presentation of experimental materials and imaginative induction followed a strict standardized procedure: (1) Trained experimenters read the unified instructions to eliminate subjective interference; (2) Text reading was limited to 10 min with a fixed speed guide (150 Chinese characters per minute), and the video was played in full without manual operation by participants; (3) After the presentation of materials, all participants in the experimental groups completed the imaginative writing task with a unified time limit of 10 min. While reading and watching the materials, participants in the self-perspective group were instructed to imagine themselves as the protagonist and write down their thoughts, feelings, and expectations for the future when facing discrimination, misunderstanding, or rejection from others (with a minimum writing length of 200 Chinese characters). Participants in the other-perspective group were instructed to imagine the plight of the protagonist with mental illness and write down the protagonist's possible thoughts, feelings, and expectations for the future in the above situations (with a minimum writing length of 200 Chinese characters). The control group did not access any of the above textual or video materials related to mental illness stigma and empathy induction; instead, they independently read the

textbook materials of the College Students' Mental Health Education course for the same duration as the self-perspective and other-perspective groups completed their reading and writing tasks (Fig. 2).

To ensure consistent manipulation intensity across groups, only the perspective-inducing sentences in the instructions were modified during the experimental operation, while all other conditions (including laboratory environment, material presentation duration, experimental task requirements, and experimenter operation) remained consistent.

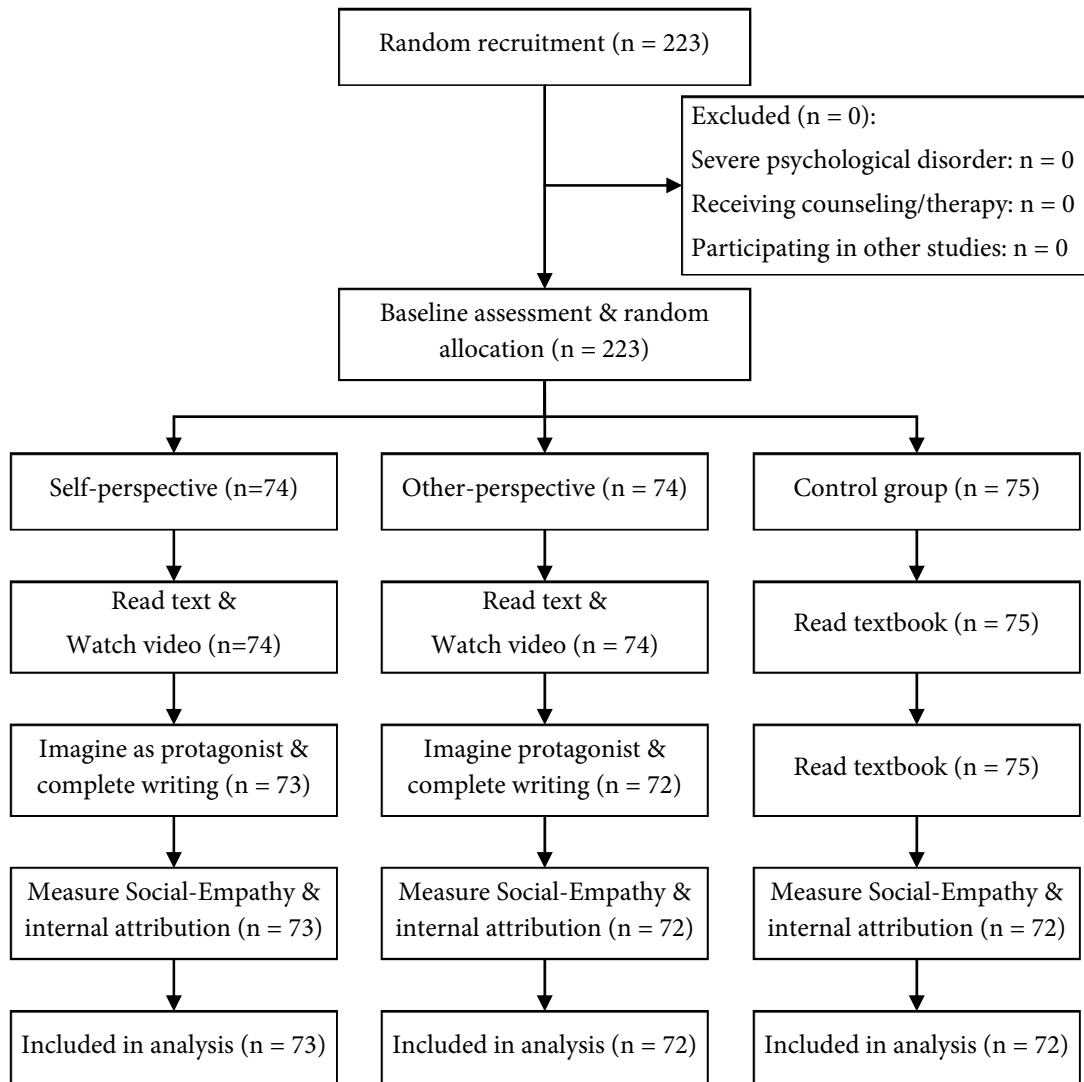


Figure 2: Flow diagram of the experimental procedure and participant allocation.

Testing the Validity of Empathy Elicitation

To comprehensively verify the effectiveness of imaginative empathy manipulation and capture the immediate state empathy post-manipulation, two complementary approaches were adopted for empathy validity testing in this study. The Social Empathy Competence Questionnaire [45] was administered immediately after the completion of the imaginative writing task (post-manipulation). Meanwhile, a content coding analysis was conducted on the imaginative writing texts, with the coding dimensions including empathic concern, identification, and transportation. Each dimension was scored on a 1–5

point scale by two independent blind raters. The inter-rater reliability was examined using the intraclass correlation coefficient (ICC), and the total score of state empathy was calculated by summing the scores of the three dimensions, which was used to reflect the intensity of immediate state empathy induced by the manipulation.

3.1.3 Statistical Analysis

General statistical procedures followed Study 1. Specific tests included chi-square test, one-way ANOVA, LSD post-hoc comparisons, repeated-measures ANOVA, and simple effect analysis.

3.2 Results

3.2.1 Imaginative Empathy Validity Test

The validity of the material reading and imagery task was verified across the self-perspective group, other-perspective group, and control group. A chi-square test was performed on the standard scores of social empathy ability, yielding a significant result ($\chi^2(2) = 11.64, p = 0.003$, Cramer's $V = 0.23$, 95% CI [0.12, 0.36]). Specifically, the proportion of participants with high social empathy was significantly higher in the self-perspective group (64.44%) than in the other-perspective group (40.96%) and control group (40.91%). In contrast, the proportion of participants with low social empathy was significantly lower in the self-perspective group (35.56%) than in the other-perspective group (59.04%) and control group (59.09%). A one-way ANOVA was conducted on the total score of social empathy ability, and the results showed a significant main effect of group ($F(2, 214) = 7.36, p < 0.001$, partial $\eta^2 = 0.06$). LSD post-hoc multiple comparisons revealed that the empathy arousal level of the self-perspective group (Mean = 12.69, SD = 4.04) was significantly higher than that of the other-perspective group (Mean = 10.95, SD = 3.70, $p = 0.003$) and the control group (Mean = 10.27, SD = 3.73, $p < 0.001$). The inter-rater reliability analysis showed that the ICC values of empathic concern, identification, transportation, and the total score of state empathy were 0.86, 0.88, 0.87, and 0.94, respectively, indicating good coding reliability. Independent samples t -test results indicated that the self-perspective group scored significantly higher than the other-perspective group in all dimensions of state empathy: empathic concern (self-perspective: Mean = 4.08, SD = 0.71; other-perspective: Mean = 2.97, SD = 0.81, $t = 8.77, p < 0.001$, Cohen's $d = 1.46$), identification (self-perspective: Mean = 4.27, SD = 0.59; other-perspective: Mean = 2.71, SD = 0.78, $t = 13.51, p < 0.001$, Cohen's $d = 2.24$), transportation (self-perspective: Mean = 3.97, SD = 0.63; other-perspective: Mean = 2.82, SD = 0.88, $t = 9.03, p < 0.001$, Cohen's $d = 1.50$), and the total score of state empathy (self-perspective: Mean = 12.32, SD = 1.10; other-perspective: Mean = 8.50, SD = 1.34, $t = 18.77, p < 0.001$, Cohen's $d = 3.12$). These results indicate that the experimental manipulation of imaginative empathy was effective, and the self-perspective induction elicited a significantly higher level of empathy arousal than the other-perspective induction. Collectively, the convergent results from the Social Empathy Competence Questionnaire and content coding of imaginative writing texts provide supportive evidence for the effectiveness of the imaginative empathy manipulation, verifying that the experimental induction successfully altered participants' state empathy levels as intended.

3.2.2 Imaginative Empathy and Situation Interact to Influence Internal Attribution Tendency

A repeated-measures ANOVA (Table 4) was conducted, with effect size magnitudes interpreted using conventional benchmarks for partial η^2 : small (0.01–0.059), medium (0.06–0.137), and large (≥ 0.138) [47]. The analysis yielded significant main effects of measurement time ($F(1, 214) = 24.49, p < 0.001$, partial $\eta^2 = 0.10$), situation ($F(1, 214) = 25.85, p < 0.001$, partial $\eta^2 = 0.11$), and group ($F(2, 214) = 3.76, p < 0.05$, partial $\eta^2 = 0.03$). Significant two-way interaction effects were observed for measurement time $\times$ group

($F(2, 214) = 4.73$, $p < 0.01$, partial $\eta^2 = 0.04$), measurement time $\times$ situation ($F(1, 214) = 63.36$, $p < 0.001$, partial $\eta^2 = 0.23$), and situation $\times$ group ($F(2, 214) = 13.33$, $p < 0.001$, partial $\eta^2 = 0.11$). Additionally, the three-way interaction of measurement time $\times$ situation $\times$ group was significant ($F(2, 214) = 11.93$, $p < 0.001$, partial $\eta^2 = 0.10$).

Table 4: Repeated-measures ANOVA results of internal attribution tendency.

Variables	SS	df	MS	F	p	Partial η^2
measurement time	59.03	1	59.03	24.49	<0.001	0.10
situation	86.35	1	86.35	25.85	<0.001	0.11
group	33.97	2	16.98	3.76	0.025	0.03
measurement time $\times$ group	22.78	2	11.39	4.73	0.009	0.04
measurement time $\times$ situation	134.56	1	134.56	63.36	<0.001	0.23
situation $\times$ group	89.03	2	44.52	13.33	<0.001	0.11
measurement time $\times$ situation $\times$ group	50.67	2	25.34	11.93	<0.001	0.10

Note: SS = Sum of Squares, df = degrees of freedom, MS = Mean Square, partial η^2 = partial eta-squared.

Simple effects analyses indicated that in positive situations (Fig. 3A), only the self-perspective group exhibited a significant increase in internal attribution tendency from pre-test to post-test ($F(1, 214) = 8.70$, $p < 0.01$, partial $\eta^2 = 0.04$); no significant changes were observed in the other-perspective ($F(1, 214) = 0.05$, $p = 0.816$, partial $\eta^2 = 0.00$) or control group ($F(1, 214) = 0.02$, $p = 0.898$, partial $\eta^2 = 0.00$). In negative situations (Fig. 3B), both the self-perspective ($F(1, 214) = 109.63$, $p < 0.001$, partial $\eta^2 = 0.34$) and other-perspective groups ($F(1, 214) = 87.97$, $p < 0.001$, partial $\eta^2 = 0.29$) showed significant reductions in internal attribution tendency post-test, the control group had no significant pre-post change ($F(1, 214) = 0.23$, $p = 0.630$, partial $\eta^2 = 0.00$).

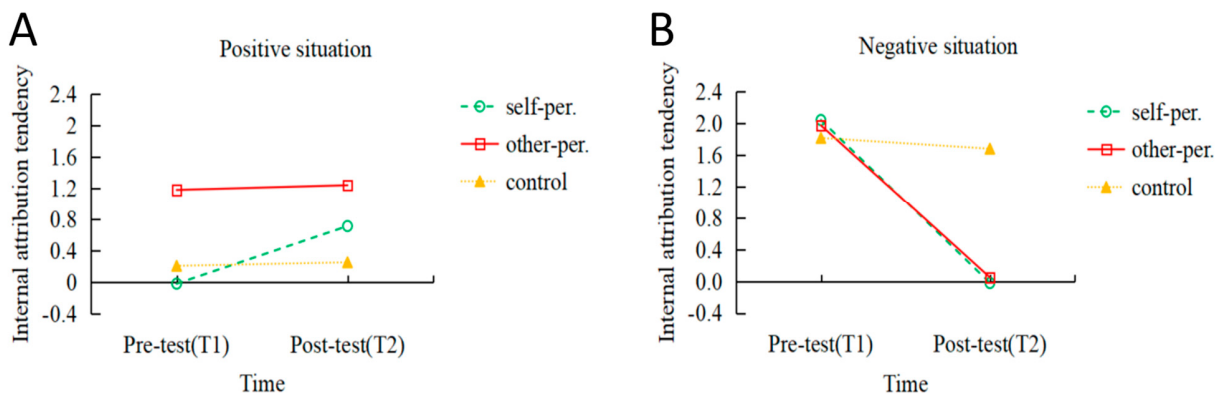


Figure 3: Simple slope plots of the interaction between Time and Condition on internal attribution tendency. (A) Positive situation; (B) Negative situation.

3.3 Discussion

In Study 2, self-perspective-based imaginative empathy significantly increased participants' internal attribution tendency in positive situations and substantially reduced it in negative situations. According to the empathy-altruism hypothesis, the affective empathy elicited by self-perspective induction evokes individuals' altruistic motivation and caring tendency, reduces the "cognitive alienation" toward patients, and further decreases the implicit attribution bias based on prejudice [50]. Attribution theory holds that individuals' causal attribution patterns toward the behaviors of individuals with mental illness

directly influence their stigmatizing attitudes. Self-perspective imaginative empathy leads participants to recognize that patients' positive behaviors result from their own abilities or characteristics (increased internal attribution, reduced stigma), and no longer simply attribute patients' negative behaviors to their own traits, but to external situational factors such as social discrimination and lack of support (reduced internal attribution, reduced stigma) [18,24]. Through the process of "identity substitution", self-perspective imaginative empathy enables participants to simulate and experience the real situation and psychological state of patients. On the one hand, it strengthens internal attribution for patients' positive behaviors (acknowledging their subjective initiative); on the other hand, it promotes situational attribution for negative behaviors (perceiving their external dilemmas). Such reconstruction of attribution patterns directly weakens the stereotypical cognition that "behavioral problems are caused by patients' own defective traits". This finding suggests that self-perspective imaginative empathy is a practical and effective approach to improving stigma-related attitudes, which fully supports Hypothesis 1. Additionally, other-perspective-based imaginative empathy reduced internal attribution tendency in negative situations but yielded no change in positive situations. This pattern may partly stem from participants' initially high internal attribution tendency in positive situations (Fig. 3A), indicating that their baseline attitudes toward individuals with mental illness were already positive, leaving limited room for further improvement during the experiment. Furthermore, other-perspective imaginative empathy reflects a perceptual/cognitive state [24], which may hinder the development of emotional empathy with the protagonists in the materials.

Collectively, these results highlight self-perspective-based imaginative empathy as a promising intervention strategy for challenging mental illness stigma.

4 Study 3: The Moderating Role of Gender in Imaginative Empathy and Stigma

4.1 Methods

4.1.1 Participants and Design

Preliminary calculations conducted using G*Power 3.1 indicated that a minimum of 54 participants would be required for the between-subjects repeated measures ANOVA to apply to this experiment (*Effect size* = 0.25, α = 0.05, *Power* = 0.95). The study utilized a 2 (time: pre-test vs. post-test, within-subjects) $\times$ 2 (situation: positive vs. negative, within-subjects) $\times$ 2 (group: male vs. female, between-subjects) experimental design. Participants were recruited through the campus information platform, and the final eligible participants were determined via a random selection method with the gender ratio controlled at approximately 1:1, covering multiple academic fields including Education, Mathematics and Applied Mathematics, and Broadcasting and Hosting. The inclusion criteria were as follows: aged 18 years or above, self-reported no history of severe psychological or mental disorders, and willing to participate in the study with experimental intervention and sign the informed consent form. The exclusion criteria included currently receiving psychological counseling or therapy, concurrently participating in other research projects, and being unable to understand the task requirements. A random sample of 71 undergraduate students was selected for the experiment (35 males, 36 females), aged between 18 and 22 years old (Mean = 19.96, SD = 0.99). No significant differences were found between the two groups in terms of age ($t = 1.62$, $p = 0.109$, Cohen's $d = 0.44$), pre-test scores of internal attribution tendency in positive situations ($t = 1.98$, $p = 0.051$, Cohen's $d = 0.44$), and pre-test scores of internal attribution tendency in negative situations ($t = 1.95$, $p = 0.055$, Cohen's $d = 0.36$).

4.1.2 Procedures

Following the registration of information by the participants, including gender and age, the Stigma Attitude Pre-test (T1) was conducted. The Imagery Paradigm was then employed to induce empathy, after

which the participants completed the Stigma Attitude Post-test (T2). The specific experimental materials and procedure were identical to those utilized in Study 2.

4.1.3 Statistical Analysis

General statistical procedures followed in Study 1. Specific tests included independent-samples *t*-test, repeated-measures ANOVA, and simple effect analysis.

4.2 Results

4.2.1 Moderation of Gender on Internal Attribution Tendency in Positive Situations

A repeated-measures ANOVA yielded a significant main effect of measurement time ($F(1, 69) = 5.88$, $p = 0.018$, partial $\eta^2 = 0.08$), a marginally significant main effect of gender ($F(1, 69) = 3.76$, $p = 0.056$, partial $\eta^2 = 0.05$), and a significant measurement time $\times$ gender interaction effect ($F(1, 69) = 4.77$, $p = 0.032$, partial $\eta^2 = 0.07$). Simple effects analyses (Fig. 4A) indicated that male participants exhibited a significant increase in internal attribution tendency at post-test ($F(1, 69) = 10.47$, $p = 0.002$, partial $\eta^2 = 0.13$), whereas no significant pre-test to post-test change was observed among female participants ($F(1, 69) = 0.03$, $p = 0.864$, partial $\eta^2 = 0.00$).

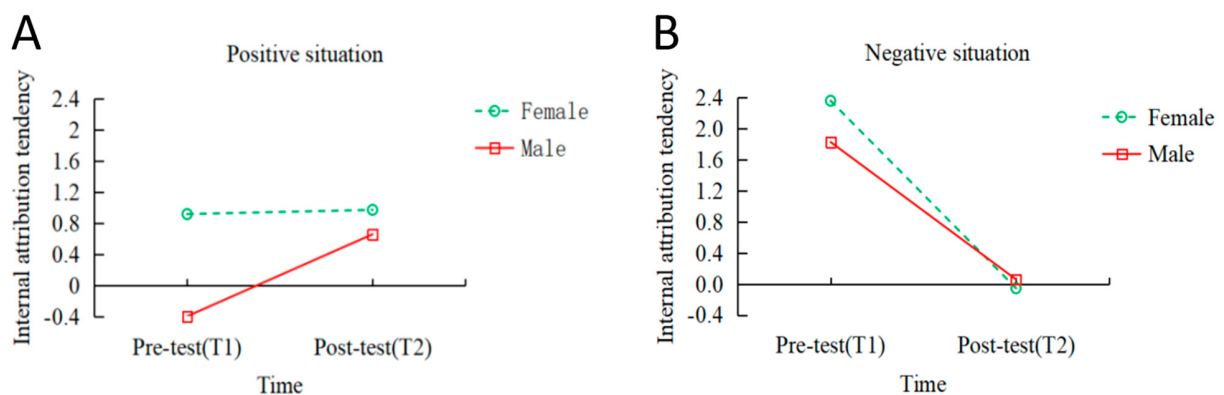


Figure 4: Simple slope plots of the interaction between Time and Gender on internal attribution tendency. (A) Positive situation; (B) Negative situation.

4.2.2 Moderation of Gender on Internal Attribution Tendency in Negative Situations

A repeated-measures ANOVA revealed a significant main effect of measurement time ($F(1, 69) = 86.29$, $p < 0.001$, partial $\eta^2 = 0.56$). Neither the main effect of gender ($F(1, 69) = 0.60$, $p = 0.440$, partial $\eta^2 = 0.01$) nor the measurement time $\times$ gender interaction effect ($F(1, 69) = 2.05$, $p = 0.157$, partial $\eta^2 = 0.03$) reached statistical significance (Fig. 4B).

4.3 Discussion

Study 3 employed a pretest-posttest experimental design to provide preliminary convergent evidence for the gender moderation effect observed in Study 1. The main effect of measurement time was significant in both positive and negative situations, with participants' internal attribution tendency shifting in a favorable direction. This finding indicates that corroborating the intervention efficacy demonstrated in Study 2. Additionally, the gender $\times$ measurement time interaction effect was significant in positive situations, which is consistent with Study 1's finding of gender moderation. While this provides preliminary convergent

support for Hypothesis 2, and the relatively small sample size ($n = 71$) and resultant limited statistical power mean this finding should be interpreted as suggestive rather than definitive.

5 General Discussion

Exposure—a stigma-reduction intervention involving direct exposure to individuals with mental illness and observing of their experiences [51]. A subsequent meta-analysis of Chinese studies revealed that non-contact interventions yielded significantly stronger effects than contact-based approaches (Hedges's $g = 0.89$ vs. 0.02 , $p < 0.001$) [18]. Notably, the effectiveness of contact-based interventions is moderated by participants' prior contact experiences with individuals with mental illness [50]. In Chinese and other Asian cultures, mental illness is often attributed to personal weakness or superstitious beliefs [52], resulting in limited daily exposure to affected individuals; poor-quality contact further compromises stigma-reduction outcomes [28].

Imaginative empathy is therefore proposed as a culturally suitable antecedent psychological process for addressing mental illness stigma in Chinese contexts, with attribution patterns serving as a critical intervening cognitive mechanism that links empathy induction to changes in stigmatizing cognitions. Study 2 demonstrated that guiding self-perspective-based imaginative empathy altered participants' internal attribution tendencies toward the behaviors of individuals with mental illness, and these attributional shifts occurred in a direction consistent with reduced implicit mental illness stigma; Study 3 further found that this empathy-attribution relationship was moderated by gender, with male participants exhibiting more pronounced attributional changes in positive situational contexts. Imagination, perception, and memory share common neurobiological substrates [29] and can elicit emotional-affective responses analogous to real-life scenarios [30]. Thus, imaginative empathy functions as a targeted induction strategy to activate emotional and cognitive empathic responses in participants, which in turn shapes their causal attribution patterns toward individuals with mental illness. The attributional changes observed in this study represent an implicit cognitive proxy for stigma reduction, aligned with the SEB framework and attribution theory's conceptualization of stigma as rooted in biased causal reasoning about marginalized groups [9,12].

From a social cognitive theory perspective, self-perspective imaginative empathy acts as a cognitive restructuring intervention [53] that targets empathic responding first (rather than stigma or attribution directly). By guiding participants to process information from the perspective of individuals with mental illness, it disrupts automatic stereotypical association (e.g., mental illness = danger/incompetence) and fosters situationally grounded cognitive processing of this group's behaviors. Integrating attribution theory and the empathy-altruism hypothesis, the causal chain of effects is: (1) self-perspective imaginative empathy induction elicits heightened emotional empathy (a core tenet of the empathy-altruism hypothesis), motivating participants to re-evaluate pre-existing causal reasoning about individuals with mental illness; (2) this re-evaluation leads to adaptive attribution shifts—increased internal attribution for positive behaviors and decreased internal attribution for negative behaviors of individuals with mental illness; (3) these attributional shifts represent a change in implicit cognitive responding that is consistent with reduced mental illness stigma, as they counteract the core biased attributions (e.g., blaming negative behaviors on stable personal deficits) that define stigmatizing cognitions toward this group [12,14]. Importantly, this chain represents a plausible causal pathway supported by the study's findings, rather than a definitive, direct causal link between empathy induction and stigma reduction itself.

Given that negative behavioral scenarios are more representative of real-world encounters with mental illness stigma, the findings derived from such contexts possess greater ecological validity and practical implications. Although the self-perspective imaginative empathy intervention produced a statistically

significant and large effect size in altering internal attribution tendencies in these realistic negative contexts (partial $\eta^2 = 0.34$)—a shift consistent with reduced stigma—it did not demonstrate absolute unique superiority in driving these attributional changes. The other-perspective empathy intervention also yielded a significant shift in internal attribution tendencies for negative behaviors (partial $\eta^2 = 0.29$), and no significant gender moderation effect was observed in the empathy-attribution relationship for negative situational contexts. Accordingly, practical anti-stigma interventions may adopt a context-specific strategic approach that targets empathic induction to shape adaptive attribution patterns: prioritizing self-perspective induction in positive behavioral scenarios to foster increased internal attribution, while integrating both self- and other-perspective strategies in negative scenarios to enhance the magnitude of adaptive attributional shifts. For instance, participants can first be guided to adopt an other-perspective to observe and imagine the external circumstances, social dilemmas, and psychological distress experienced by individuals with mental illness, thereby fostering cognitive understanding of their predicament. Concurrently, participants can be instructed to shift to a self-perspective to imagine and experience the fear, hurt, and helplessness they would feel if subjected to the same stigmatizing situations, which in turn strengthens emotional engagement with the target group. This combined perspective strategy compensates for the inherent limitations of a single-perspective approach and further enhances the stability and ecological validity of anti-stigma interventions in real-world settings.

In the present study, PI–PE (Positive situational Internal minus External attribution) and NI–NE (Negative situational Internal minus External attribution) were employed as implicit attribution-based proxies for stigmatizing attitudes toward individuals with mental illness, with the core rationale rooted in the Stereotypic Explanatory Bias (SEB) framework and attribution theory. Critically, this approach does not constitute a direct measurement of mental illness stigma in its broader sense. The interpretation of these attributional proxies is subject to alternative plausible explanations, and thus their scope of application requires explicit clarification. First, an increase in internal attribution tendency in positive situational contexts, while operationalized as a marker of reduced stigma in our SEB-based paradigm, may also reflect an elevation in participants' perceived agency, optimism, or competence judgments of individuals with mental illness. Notably, this elevation in positive cognitive appraisals is inherently synergistic with stigma reduction, as both represent meaningful positive shifts in participants' implicit cognitive attitudes toward individuals with mental illness [22], and thus this alternative interpretation does not contradict our core finding of reduced stigmatizing cognitions. Second, a decrease in internal attribution tendency in negative situational contexts is likely associated with participants' empathic understanding: empathy enhances participants' ability to perceive the external adversities faced by individuals with mental illness, which in turn increases the propensity for external attributions of their negative behaviors. This cognitive-emotional process constitutes a key psychological mechanism through which empathy ameliorates stigmatizing cognitions, rather than merely reflecting a simple shift in attributional bias [26]. Third, as implicit cognitive proxies, these attributional metrics only reflect participants' causal explanatory biases toward the behaviors of individuals with mental illness, and cannot be equated entirely with explicit stigmatizing attitudes (e.g., social distance, discriminatory intent) or affective stigma (e.g., fear, aversion). Future research should adopt a multi-metric approach that combines these implicit attribution proxies with established direct stigma scales and behavioral experiments to further verify the convergent validity between attributional tendencies and mental illness stigma in its broader sense.

Post-pubertal gender differences in gonadal function have been linked to variations in empathic responses in prior research [41,42], with oxytocin proposed to facilitate empathy processing in females and testosterone suggested to dampen empathic responses in males. These sex-related hormonal regulatory

mechanisms are a theoretical hypothesis for the observed gender differences in the present study, such mechanisms could potentially underpin the relatively lower trait empathy tendencies in male participants, and thus the more pronounced improvement in their stigmatizing attitudes following the same imaginative empathy intervention [44]. Notably, gender differences in empathic responses and stigma perception are not solely determined by biological factors, but are also shaped by sociocultural processes such as gender role socialization and internalization of social norms (see Section 2.3 Discussion for details). Specifically, social expectations that females should be more nurturing and emotionally sensitive, while males should be more independent and emotionally restrained, further amplify the gender differences in empathy and stigma attitudes observed in the present study. However, as the current research did not include direct measurements of hormonal levels, psychophysiological indicators, or sociocultural variables (e.g., gender role identification), this dual-pathway interpretive framework (biological and sociocultural) remains untested here. Future research is needed to empirically verify this integrated hypothesis by integrating hormonal assays, neurophysiological measurements, and assessments of gender role beliefs into the experimental design.

6 Limitations and Implications

First, given the absence of a universally standardized measurement tool for empathy [28], the Social Empathy Competence Questionnaire utilized in this study fails to capture the full connotation of the empathy construct. Second, self-perspective-based imaginative empathy may introduce potential biases: without a clear distinction between self and others, individuals may overfocus on their own emotional experiences, developing a “comparative luck” mindset that neglects the actual needs of people with mental illness and overestimates their own adaptive responses to adversity. Such over-imagining of personal reactions may further inhibit empathic concern for the target group [20]. Future research should therefore develop methodological strategies to mitigate these biases during the imaginative empathy intervention, enhance participants’ awareness of the differences between individuals with mental illness and themselves, and further refine the intervention protocol. Third, the experimental design of Study 3 has certain methodological limitations. Although Study 2 has verified the effectiveness of the imaginative empathy intervention using a controlled design, Study 3 adopted a single-group pretest-posttest design without an independent control or placebo group to focus on the gender moderation effect. As a result, the observed changes from pretest to posttest cannot completely exclude the influence of the Hawthorne effect, social desirability, or repeated testing, which may reduce the internal validity of the findings. Fourth, the present study only assessed the immediate effect of the intervention and did not include long-term follow-up measurements. Therefore, the persistence and stability of the intervention effect over time remain unclear. The lack of follow-up tests was mainly constrained by the research period, potential participant attrition caused by holidays or graduation, and limited research resources. Future studies are strongly recommended to adopt a longitudinal design with multiple follow-up time points to verify the long-term effect of the imaginative empathy intervention. Fifth, the present study was conducted exclusively among Chinese undergraduate students aged 18–22 years, and the sample is highly homogeneous in terms of cultural background, educational level, and age group; in addition, the gender ratio was imbalanced in partial sub-samples. This results in limited generalizability of the research findings beyond Chinese undergraduates, and the conclusions cannot be directly applied to other cultural contexts, age groups, or social groups with different educational backgrounds. Sixth, the sub-sample size of Study 3 was relatively small, and the variability of the internal attribution tendency index was not fully considered in the sample size calculation, which may have led to insufficient statistical power. Future research should determine a larger and more reasonable sample size for different gender

subgroups to ensure sufficient statistical power for small to medium effect sizes of gender moderation effects, reduce the risk of statistical errors, and improve the reliability of research results. Future studies are therefore encouraged to adopt more inclusive sampling strategies covering a wider range of ages, educational levels, geographic regions, and developmental backgrounds, and ensure balanced gender representation in sampling; the research samples should also be expanded to non-college student groups and cross-cultural populations. Comparative investigations across heterogeneous groups would help clarify the differential effects and boundary conditions of imaginative empathy interventions in diverse populations, and further verify the cross-cultural adaptability and general applicability of the intervention strategy. Additionally, the attribution coding in this study was based on manual judgment by trained coders despite strict reliability verification; future research can combine natural language processing (NLP) technology to realize automated coding of free-text responses, further reducing subjective coding biases. Seventh, the manipulation check for imaginative empathy has an inherent endogeneity issue: the blind content coding of empathic concern, identification, and transportation was derived from the imaginative writing texts produced by participants during the intervention manipulation itself, which makes this coding measure closely intertwined with the manipulation and not a fully independent assessment of state empathy. Although we complemented this with the Social Empathy Competence Questionnaire—a psychometric scale administered post-manipulation—to objectively measure empathy levels, the state-empathy evidence from the combined approaches remains supportive rather than definitive for verifying the manipulation effect. Future research should develop more independent manipulation check measures, such as administering a separate, non-task-related state empathy scale or conducting post-experiment semi-structured interviews, to avoid the confounding of manipulation tasks and validity assessment, and further improve the methodological rigor of empathy induction verification.

The findings imply that to effectively reduce Chinese undergraduates' stigmatizing attitudes toward individuals with mental illness, it is imperative to design culturally tailored narratives and scenarios that align with the cognitive characteristics of Chinese students. These interventions should guide students to envision themselves in the position of individuals with mental illness or psychological distress, thereby fostering deep emotional empathy and reducing stigmatizing cognitions.

7 Conclusions

Guided by the cultural characteristics of mental illness stigma in China, this study adapted content from mainstream news media reports as experimental stimuli, employed the stereotypic explanatory bias (SEB) paradigm, and developed a behavioral attribution task that is consistent with Chinese cultural norms. The results validated that self-perspective imaginative empathy is an effective intervention strategy for mitigating attribution-based implicit stigmatizing cognitions (a core component of mental illness stigma) among Chinese undergraduate students—an important proxy for overall stigma reduction in the Chinese cultural context. Notably, the efficacy of this intervention varies by gender: it exerts a more significant improvement effect on male students' attributional biases toward stigmatizing cognitions in positive behavioral scenarios, while the moderating effect of gender on attributional shifts is not significant in negative behavioral scenarios.

Acknowledgement: Not applicable.

Funding Statement: This research was funded by the Anhui Provincial Department of Education Project (Project Name: *Construction and Intervention Research on Prediction Models of College Students' Social Exclusion*; Grant Number: 2025AHGXSK30539), and the Anhui Provincial Department of Education Project (Project Name: *Study*

on the Influencing Factors and Mechanism of Implicit Stigma of Mental Illness Among College Students; Grant Number: SK2021A0902).

Author Contributions: The authors confirm contribution to the paper as follows: Conceptualization, Siliang Yang; methodology, Siliang Yang; software, Siliang Yang; validation, Siliang Yang and Gan Jin; formal analysis, Siliang Yang; investigation, Siliang Yang; resources, Siliang Yang; data curation, Siliang Yang; writing—original draft preparation, Siliang Yang; writing—review and editing, Siliang Yang and Gan Jin; visualization, Siliang Yang; supervision, Siliang Yang; project administration, Siliang Yang; funding acquisition, Siliang Yang. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The data that support the findings of this study, as well as the experimental materials (e.g., the expert-verified textual and video materials), are available from the Corresponding Author upon reasonable request.

Ethics Approval: All procedures involving human participants across the three sub-studies complied with the guidelines of the institutional research ethics committee and the Declaration of Helsinki. This study was reviewed and approved by the Academic Ethics Committee of Anhui Normal University, China (Approval No.: AHNU-ET2022002) on 17 February 2022. Prior to participating in any data collection or experimental procedures, all participants were fully informed of the research purpose, procedures, and their right to withdraw from the study at any time without incurring any penalty. The participants signed the informed consent form.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Li SA, Sakata N, Donn G. A scoping review of mental health literacy in performing and creative artists: Identifying current gaps and future directions. *Front Psychol.* 2025;16:1329029. [[CrossRef](#)].
2. Huang Y, Wang Y, Wang H, Liu Z, Yu X, Yan J, et al. Prevalence of mental disorders in China: A cross-sectional epidemiological study. *Lancet Psychiatry.* 2019;6(3):211–24. [[CrossRef](#)].
3. Yu G, Huang X. A comparison of the rates of prevalence of students' mental health problems: Evidence based on meta-analysis. *Educ Res.* 2023;44(06):105–21. (In Chinese).
4. Ning X, Wong JP, Huang S, Fu Y, Gong X, Zhang L, et al. Chinese university students' perspectives on help-seeking and mental health counseling. *Int J Environ Res Public Health.* 2022;19(14):8259. [[CrossRef](#)].
5. Phillips MR, Zhang J, Shi Q, Song Z, Ding Z, Pang S, et al. Prevalence, treatment, and associated disability of mental disorders in four provinces in China during 2001–05: An epidemiological survey. *Lancet.* 2009;373(9680):2041–53. [[CrossRef](#)].
6. Singh S, Melendez K, Sezginis N. Examining the effect of discrimination and stigma on utilization of mental health services among college students. *J Am Coll Health.* 2023;71(8):2398–405. [[CrossRef](#)].
7. Louzán R, Villarroja López A. Impact of social desirability bias on the CoPsoQ-ISTAS 21 psychosocial risk assessment questionnaire. *Work.* 2024;79(1):241–51. [[CrossRef](#)].
8. Hastie R. Causes and effects of causal attribution. *J Pers Soc Psychol.* 1984;46(1):44–56. [[CrossRef](#)].
9. Sekaquaptewa D, Espinoza P, Thompson M, Vargas P, von Hippel W. Stereotypic explanatory bias: Implicit stereotyping as a predictor of discrimination. *J Exp Soc Psychol.* 2003;39(1):75–82. [[CrossRef](#)].
10. Kulik JA. Confirmatory attribution and the perpetuation of social beliefs. *J Pers Soc Psychol.* 1983;44(6):1171–81. [[CrossRef](#)].
11. Wen FF, Gao J, Zuo B, He SF, Lin YT, Wang J, et al. Examining children's implicit gender stereotyping using the stereotypic explanatory bias paradigm. *Int J Psychol.* 2023;58:212.
12. Kågström A, Guerrero Z, Aliev AA, Tomášková H, Rüscher N, Ouali U, et al. Mental health stigma and its consequences: A systematic scoping review of pathways to discrimination and adverse outcomes. *eClinicalMedicine.* 2025;89:103588. [[CrossRef](#)].
13. Andersen MM, Varga S, Folker AP. On the definition of stigma. *Eval Clin Pract.* 2022;28(5):847–53. [[CrossRef](#)].
14. Davis ML, Witcraft SM, Baird SO, Smits JAJ. Learning principles in CBT. In: *The science of cognitive behavioral therapy.* Amsterdam, The Netherlands: Elsevier; 2017. p. 51–76. [[CrossRef](#)].

15. Byrne SJ, Swords L, Nixon E. Depression literacy and self-reported help-giving behaviour in adolescents in Ireland. *Child Psychiatry Hum Dev.* 2026;57:543–55. [[CrossRef](#)].
16. Zeigler-Hill V, Shackelford TK. *Encyclopedia of personality and individual differences.* Cham, Switzerland: Springer International Publishing; 2018. [[CrossRef](#)].
17. Miller L, Musci R, D'Agati D, Alfes C, Beaudry MB, Swartz K, et al. Teacher mental health literacy is associated with student literacy in the adolescent depression awareness program. *Sch Ment Health.* 2019;11(2):357–63. [[CrossRef](#)].
18. Ren Z, Zhao C, Tian F, Yan Y, Li D, Zhao Z, et al. Meta-analysis of the effect of mental health literacy intervention in Chinese people. *Acta Psychol Sin.* 2020;52(4):497–512. (In Chinese). [[CrossRef](#)].
19. Xu Z, Huang F, Kösters M, Staiger T, Becker T, Thornicroft G, et al. Effectiveness of interventions to promote help-seeking for mental health problems: Systematic review and meta-analysis. *Psychol Med.* 2018;48(16):2658–67. [[CrossRef](#)].
20. Fan X. Cultural contours of depression: a socio-political and ethnographic analysis. In: *Proceedings of the 2024 6th International Conference on Literature, Art and Human Development (ICLAHD 2024); 2024 Oct 18–20; Xi'an, China.* [[CrossRef](#)].
21. Booth WA, Abuhmida M, Anyanwu F. Mental health stigma: a conundrum for healthcare practitioners in conservative communities. *Front Public Health.* 2024;12:1384521. [[CrossRef](#)].
22. Krafft J, Ferrell J, Levin ME, Twohig MP. Psychological inflexibility and stigma: A meta-analytic review. *J Context Behav Sci.* 2018;7:15–28. [[CrossRef](#)].
23. Stotland E. Exploratory investigations of empathy. In: Berkowitz L, editor. *Advances in experimental social psychology.* Vol. 4. New York, NY, USA: Academic Press; 1969. p. 271–313. [[CrossRef](#)].
24. Batson DC. *A scientific search for altruism: Do we only care about ourselves?* Oxford, NY, USA: Oxford University Press; 2018. p. 32–5. [[CrossRef](#)].
25. Branchadell V, Poy R, Ribes-Guardiola P, Segarra P, Moltó J. Psychopathic callousness and perspective taking in pain processing: an ERP study. *Social Cogn Affect Neurosci.* 2024;19(1):nsae022. [[CrossRef](#)].
26. Liao T, Huang Y, Chen X. Grandparent-grandchild intergenerational contact and attitudes toward older people: The chain mediating effects of intergroup anxiety and empathy. *Stud Psychol Behav.* 2024;22(1):86–92. (In Chinese). [[CrossRef](#)].
27. Tassinari M, Aulbach MB, Harjunen VJ, Cocco VM, Vezzali L, Jasinskaja-Lahti I. The effects of positive and negative intergroup contact in virtual reality on outgroup attitudes: testing the contact hypothesis and its mediators. *Group Process Intergroup Relat.* 2024;27(8):1773–98. [[CrossRef](#)].
28. Haraguchi K, Maeda M, Mei YX, Uchimura N. Stigma associated with schizophrenia: Cultural comparison of social distance in Japan and China. *Psychiatry Clin Neurosci.* 2009;63(2):153–60. [[CrossRef](#)].
29. Xiao GL, Liu MF. Neurocognitive mechanisms of mental imagery-based fear conditioning in individuals with social anxiety. *Behav Res Ther.* 2026;200:105013. [[CrossRef](#)].
30. Mitra S, Asthana MK. Attenuating conditioned fear using imagery-based interventions: an overview. *Neurosci Biobehav Rev.* 2024;163:105751. [[CrossRef](#)].
31. Emery LF, Hughes EK, Muise A. Self-expansion theory: Origins, current evidence, and future horizons. *Soc Personal Psych.* 2025;19(9):e70082. [[CrossRef](#)].
32. Ma MZ, Chen SX, Wang X. Collective pronouns, collective health actions: Predicting pandemic precautionary measures through online first-person plural pronoun usage across U.S. states. *Soc Sci Med.* 2024;357:117167. [[CrossRef](#)].
33. Decety J, Sommerville JA. Shared representations between self and other: A social cognitive neuroscience view. *Trends Cogn Sci.* 2003;7(12):527–33. [[CrossRef](#)].
34. Lamm C, Bukowski H, Silani G. From shared to distinct self–other representations in empathy: Evidence from neurotypical function and socio-cognitive disorders. *Philos Trans R Soc B Biol Sci.* 2016;371(1686):20150083. [[CrossRef](#)].
35. Schilbach L. Towards a second-person neuropsychiatry. *Phil Trans R Soc B.* 2016;371(1686):20150081. [[CrossRef](#)].
36. McAuliffe K, Dunham Y. Group bias in cooperative norm enforcement. *Phil Trans R Soc B.* 2016;371(1686):20150073. [[CrossRef](#)].

37. Cooke AN, Bazzini DG, Curtin LA, Emery LJ. Empathic understanding: Benefits of perspective-taking and facial mimicry instructions are mediated by self-other overlap. *Motiv Emot.* 2018;42(3):446–57. [CrossRef].
38. Lausi G. How cognitive processes shape implicit stereotypes: A literature review. *Open Res Europe.* 2025;4:263. [CrossRef].
39. Hemp CA. The association between self-identified gender, empathy, gender role perspectives, and attitudes toward undergraduates with autism [master's thesis]. Lowell, MA, USA: University of Massachusetts; 2022.
40. Chang-Arana ÁM, Surma-aho A, Hölttä-Otto K, Sams M. Under the umbrella: components of empathy in psychology and design. *Des Sci.* 2022;8:e20. [CrossRef].
41. Zhao W, Yao S, Li Q, Geng Y, Ma X, Luo L, et al. Oxytocin blurs the self-other distinction during trait judgments and reduces medial prefrontal cortex responses: Oxytocin Blurs Self-Other Distinction. *Hum Brain Mapp.* 2016;37(7):2512–27. [CrossRef].
42. van Honk J, Schutter DJ, Bos PA, Kruijt AW, Lentjes EG, Baron-Cohen S. Testosterone administration impairs cognitive empathy in women depending on second-to-fourth digit ratio. *Proc Natl Acad Sci U S A.* 2011;108(8):3448–52. [CrossRef].
43. Procyshyn TL, Dupertuys J, Bartz JA. Neuroimaging and behavioral evidence of sex-specific effects of oxytocin on human sociality. *Trends Cogn Sci.* 2024;28(10):948–61. [CrossRef].
44. Martinez LR, Xu S, Hebl M. Utilizing education and perspective taking to remediate the stigma of taking antidepressants. *Community Ment Health J.* 2018;54(4):450–9. [CrossRef].
45. Zhang Y. Study on the cultivation of interpersonal relationship intelligence of normal university students. *Nanjing J Soc Sci.* 2016;12:114–39. (In Chinese). [CrossRef].
46. Cheng Y, Yang S, Sang Q. Relieving implicit mental illness stigma against college students: A research based on stereotypic explanatory bias. *Chin J Spec Educ.* 2015;4:9–15. (In Chinese).
47. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. London, UK: Routledge; 1988. [CrossRef].
48. Li Y. A report on China's mental patients. *China Econ Wkly.* 2001;28:24–37. (In Chinese).
49. CCTV.com. News 1+1: Mental Disorders are Cured, How to Cure Social Stigma? [Video]. 2013 [cited 2026 Jan 1]. Available from: <https://tv.cctv.cn/2013/05/06/VIDE1367851679631329.shtml>. (In Chinese).
50. Batson CD, Lishner DA, Stocks EL. The empathy-altruism hypothesis. In: Schroeder DA, Graziano WG, editors. *The Oxford handbook of prosocial behavior*. Oxford, UK: Oxford University Press; 2015. p. 260–81. [CrossRef].
51. Corrigan P, Michaels PJ, Morris S. Do the effects of antistigma programs persist over time? Findings from a meta-analysis. *Psychiatr Serv.* 2015;66(5):543–6. [CrossRef].
52. Abdullah T, Brown TL. Mental illness stigma and ethnocultural beliefs, values, and norms: An integrative review. *Clin Psychol Rev.* 2011;31(6):934–48. [CrossRef].
53. Bandura A. Social cognitive theory: An agentic perspective. *Annu Rev Psychol.* 2001;52:1–26. [CrossRef].