Threat to the self at the heart of depression: Mediating role and depressogenic prism hypothesis

Céline Stinus and Sophie Berjot

Université de Reims Champagne-Ardenne, Reims, France

Abstract: Several psychological models tacitly highlight the role of self-threat (ST) as an entry point to depression. The first study (*n* = 311; 270 females, *M* age = 34.4) aims to demonstrate that ST is a core mechanism linking dysfunctional attitudes (DA) to depression. The second study (*n* = 263; 183 females, *M* age = 20.4) tests the «depressogenic prism» hypothesis, proposing that individuals with high DA appraise reality as invariably threatening for the self, regardless of tangible circumstances. In Study 1, ST was strongly and positively associated with both DA and depressive symptoms and served as a full mediator between DA and depressive symptoms. In Study 2, high-DA individuals exposed to a negative scenario reported significantly greater feelings of threat than their low-DA counterparts. Notably, high-DA individuals in a positive scenario felt as threatened as low-DA individuals in a negative one, underscoring the pervasive impact of DA on threat perception.

Keywords: Self-threat; cognitive appraisal; social threat; depressive vulnerability; dysfunctional attitudes.

Amenaza al yo en el núcleo de la depresión:
Papel mediador e hipótesis del prisma depresógeno

Resumen: Varios modelos psicológicos destacan tácitamente el papel de la amenaza al yo (ST) como punto de partida para la depresión. El primer estudio (*n* = 311; 270 mujeres, *M* edad = 34.4) busca demostrar que la ST es un mecanismo clave que vincula las actitudes disfuncionales (DA) con la depresión. El segundo estudio (*n* = 263; 183 mujeres, *M* edad = 20.4) pone a prueba la hipótesis del «prisma depresógeno» que sugiere que las personas con DA alta perciben la realidad como inherentemente amenazante para el yo. En el Estudio 1, la ST mostró una fuerte correlación tanto con DA como con los síntomas depresivos, actuando como mediador entre ambos. En el Estudio 2, quienes presentaron DA alta se sintieron más amenazados en escenarios negativos y, notablemente, tan amenazados en escenarios positivos como aquellos con DA baja en situaciones negativas, subrayando el impacto generalizado de las DA en la percepción de amenaza.

Palabras clave: Amenaza al yo; evaluación cognitiva; amenaza social; vulnerabilidad depresiva; actitudes disfuncionales.

Introduction

Much like the old Latin adage, "all roads lead to Rome", the same can be said about depression: "all depression models lead to the self". Interestingly, despite using theoretically distant models, many authors from separate fields of psychology have independently reached the same conclusion: threats to the self-seem to be at the heart of depression (Allen & Badcock, 2003; Beck, 1967; Dickerson et al., 2009; Kopala & Zuroff, 2020). From an evolutionary perspective, depression is not merely a by-product of modern society but rather an adaptive response embedded deeply within social species. As social beings, humans narrowly depend on their group for survival and vital resources (e.g., food, shelter, reproductive opportunities, support in raising offspring; Baumeister & Leary, 1995; Elimari

& Lafargue, 2020, 2023b). We therefore evolved mechanisms (1) to anticipate and detect risks of social exclusion and (2) to trigger adapted behaviors mitigating social rejection (Baumeister & Leary, 1995). Depression thus constitutes a "protective" neuropsychobiological program (Beck & Bredemeier, 2016) that drives social organisms to behave in a way that prevents social exclusion or moderates its consequences when "social survival" feels threatened (Allen & Badcock, 2003).

In the same vein, social psychologists consider depression as a social disease and argue that "*all social stressors can be conceptualized as 'identity interruptions' or threats to self-concept*" (Sharma & Sharma, 2010, p.129) and should be predictive of depressive symptoms (Dickerson et al., 2009). The identity threat concept (Berjot and Girault-Lidvan, 2009; Skinner & Brewers, 2002), the cognitive evaluation theory (Ryan & Deci, 2002), the concept of contingencies of self-esteem (Crocker, & Park, 2004), and the social rank theory (Wetherall et al., 2019), are as many frameworks illustrating the plurality of ways in which society can "*injure*" the self, and thus leads to depression.

The notion of "damaged self-concept" evoked by Sharma & Sharma (2010) is reminiscent of Beck's cognitive theory of depression (1967) which postulated that negative beliefs about the self (lovable vs. unlovable), others (accepting vs. rejecting), and the future (hopeful vs. hopeless) constitute a major vulnerability factor for depression (Alloy et al., 2000). People prone to depression view themselves and the world through the lens of their depressiogenic schemas, prompting them to feel inadequate in social situations that touch on aspects of the self. Interestingly, although Beck's cognitive triad posited that negative thoughts affect one's view of the self, others, and the future, Pietromonaco (1985) found that negative beliefs are primarily centered around the self, suggesting once again that a wounded self lies at the heart of the depressive puzzle.

The concept of self is not merely a subject of interest for modern psychologists. For millennia, philosophers such as Plato, Diogenes and Nāgārjuna, as well as spiritualities like Taoism, Buddhism, and Christianity, have all wondered about how we consistently find the self in the midst of human torment. For instance, according to the 2500-years-old Buddhist teachings, suffering (including what we call depression in modern terms) comes from the mismatch between our erroneous sense of self and the veritable nature of existence: while we evolve in a reality that is intrinsically impermanent (i.e., constantly changing moment-by-moment) and interdependent (i.e., all phenomenon occurs in dependance to each other),

we consider ourselves as an "inherent" entity (i.e., stable over time and minimally influenced by external factors, Bstan-vdzin-rgya-mtsho, 1997 ; Varela et al., 2017). This illusory self creates a continuous feeling of threat that pushes us into a losing battle against this ever-changing reality (Dambrun & Ricard, 2011). To translate those ideas into modern psychological terms, this ever-changing reality threatens the set of self-related beliefs and processes that collectively form our sense of a stable self.

This sense of an illusory self is not inherently wrong or aberrant, it is a natural by-product of the fact that we do not perceive ourselves and the world as they objectively are but rather in the light of our own biases (Beck & Bredemeier, 2016). In a recent review, Petriglieri (2011) stressed out the subjective dimension of threat to the self-arguing that "*identity threat arises from an individual's subjective appraisal of an experience as indicating potential harm to one of the individual's identities*" (p. 642). In other terms, it is not so much the threat to the self itself, but rather how we perceive and evaluate this threat, that serves as the real gateway to depression. The same event can be a source of distress for one person, yet have no significance for another: when it comes to self-threat: subjectivity is key.

Empirical findings from psychology and neuroscience underscore the importance of subjectivity, showing that the valence of any situation arises from human interpretation (Camuñas et al., 2019; Gadassi & Rafaeli, 2015; Kuiper et al., 1988; Lazarus, 1991; Ruiz-Alonso et al., 2021). Neuroscientific studies demonstrated that even the most unequivocal form of suffering (i.e., physical pain) find its reality in subjectivity : for instance, modulating perceived pain intensity via hypnosis without changing objective conditions leading to pain (i.e., water temperature leading to a burning sensation) actually modifies neural activation patterns in sensory cortices and limbic structures (Rainville et al., 1997). In the context of depressive vulnerability, sociobiological data showed that the more people feel socially evaluated, the more they secrete proinflammatory cytokines (Dickerson et al., 2009) a post-injury inflammatory response that contributes to depression onset. Numerous studies in clinical and social psychology investigating the way individual dispositions influence cognitive appraisal found that people high in neuroticism tend to appraise daily events as negative and display greater emotional reactivity following negative appraisal (e.g., Espejo et al., 2012; Gunthert et al., 1999). Taken together, these findings suggest that interindividual differences in threat appraisal determine depressive vulnerability more than objective threat itself.

According to Beck's cognitive theory, those distorted interpretations of reality are the product of negative schemas and dysfunctional attitudes (DA) about oneself, others, and the futures (Beck, 1967). Schemas are cognitive structures that determine how a stimulus is screened, coded, and evaluated (Disner et al., 2011). As a result, they act as a filter that determines the tinge that any experience has. When schemas are depressogenic they lead to information-processing biases (Scher et al., 2005) and constitute a "depressogenic prism", a sort of grey-tinted glasses through which reality is assessed.

Based upon the empirical and theoretical elements presented above, we argue that DA constitute a depressogenic prism whereby people appraise social situations as inherently threatening to the self, regardless of actual threat. In turn, self-threat (ST) triggers depressive symptoms (Cohen et al., 2012). Hence, ST is, in a way, the engendering agent behind depressive symptoms. The aim of the present article is straightforward: to highlight how threats to the self – experiences that undermine, challenge, or jeopardize a person's sense of self, Berjot & Bourguignon, 2023; Dickerson et al., 2009) lie at the core of depression. To this end, we will first investigate the role of self-threat (ST) in the well-established connection between negative schemas, operationalized as dysfunctional attitudes, and depressive symptoms. Next, we will introduce and test the "depressogenic prism" hypothesis, which clarifies the link between dysfunctional attitudes and self-threat. Our research design and hypotheses were built upon the preliminary but essential findings detailed in Appendix A.

Method

Participants

In Study 1, a total of 311 participants completed the survey (270 women, 41 men). Mean age was 34.4 years (SD = 12.9); most were between 30 and 59 years old (73.95%). According to the Monte Carlo power analysis, a sample of 157 participants, accounting for a 30% dropout rate in incomplete questionnaires, ensured the detection of a significant indirect effect at 80%. Based on the results from Şenormancı et al. (2014), we assumed β = 0.29 for the path a (dysfunctional attitude $\rightarrow$ self-threat), β = 0.29 for the path b (self-threat $\rightarrow$ depressive symptoms), and β = 0. 06 for the path c' (dysfunctional attitudes $\rightarrow$ depressive symptoms).

In Study 2, our sample comprised 183 women (69.6%), 67 men (25.5%) and 13 individuals who identified as non-binary (4.9%) with a total of 263 students from the French University of Reims (France).

Mean age was 20.41 years (SD = 2.13). Given that our scenarios referred to a context of succeeding (versus failing) end-of-semester's exams, we restricted the sample to university students only.

Measures

Dysfunctional Attitudes Scale Short-form (DAS-9; Beevers et al., 2007). The French version validated by Bouvard et al. (1994) was used in Study 1 to assess depressogenic schemas. Since DA are built based on negative schemas (Weissman & Beck, 1978), negative schemas have been frequently operationalized at the level of DA and assessed by means of the Dysfunctional Attitudes Scale (Weissman & Beck, 1978). The short form is unidimensional and consists of 9 items rated on a four-point Likert scale, ranging from 1 (totally disagree) to 4 (totally agree), resulting in a total score that ranges from 9 to 36. Higher scores reflect increased levels of dysfunctional attitudes. In Study 1, Cronbach's alpha was acceptable (α = 0.78), comparable to the original validation where Cronbach's alpha was 0.83 (Beevers et al., 2007).

Threat Appraisal Subscale of the Cognitive Appraisal Scale (Skinner & Brewer, 2002). In Study 1, we used the French version validated by Berjot and Girault-Lidvan (2009). This subscale assesses the tendency to feel threatened in social contexts. More specifically, it measures the degree of perceived threat (e.g., negative evaluation, overt rejection) to one's self-esteem and identity. The 10-items of the threat appraisal subscale are rated along a six-point Likert scale ranging from 1 (totally disagree) to 6 (totally agree), resulting in a total score range of 10 to 60. Higher scores indicate a stronger tendency to perceive situations as self-threatening. In Study 1, the scale demonstrated excellent reliability (α = 0.94), comparable to the original French validation with a Cronbach's alpha of 0.88 (Berjot & Girault-Lidvan, 2009).

Depression subscale of the Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983). In Study 1, we used the French version validated by Bocéréan and Dupret (2014). The depression subscale comprises 7 items, each scored from 0 to 3, resulting in a total score range of 0 to 21 that reflects the presence and severity of depressive symptoms. Zigmond and Snaith (1983), the original authors of the scale, established cut-off points: scores of 0 to 7 indicate no symptoms, 8 to 10 suggest moderate symptoms, and 11 or higher reflect significant depressive symptomatology. However, Bjelland et al. (2002) emphasized that a threshold score of 8 is crucial to ensure that individuals with depressive disorders are not overlooked. In Study 1, Cronbach's alpha was good (α = 0.88), comparable to the original

French validation where Cronbach's alpha was 0.78 (Bocéréan & Dupret, 2014).

Dysfunctional attitudes scale short-form (DAS-9; Beevers et al., 2007). As in Study 1, the French version validated by Bouvard et al. (1994) was used in Study 2 to assess depressogenic schemas, with a satisfactory Cronbach's alpha ($\alpha = 0.77$). The 9-item scale, rated on a 4-point Likert scale (ranging from 1 to 4), assesses the level of dysfunctional attitudes, with higher scores indicating greater levels of these attitudes.

Negative affect subscale of the Positive and Negative Affect Schedule (Watson et al., 1988). In Study 2, we used the French version validated by Gaudreau et al. (2006) to assess the extent of negative emotions experienced by participants. Students responded to the 10-item subscale, indicating how they would feel emotionally in the previously presented scenario, using a 5-point scale ranging from 1 to 5. Higher scores indicate a higher level of negative affect. In Study 2, the reliability of the scale was good ($\alpha = 0.90$), similar to the original French version, with values ranging from 0.80 to 0.84 (Gaudreau et al., 2006).

Threat to Personal Identity subscale from the Primary Appraisal of Identity Threat scale (Berjot et al., 2012). In Study 2, we used the original French version of the 6-item subscale to assess self-threat appraisal in the assigned situation. This subscale measures how individuals perceive their sense of self as potentially questioned (e.g., "I felt I was being reassessed as a person"), threatened (e.g., "I experienced this situation as a threat to my personal identity"), or diminished (e.g., "I felt as though I was regarded as insignificant") within a given context. Participants rated their responses on a 5-point Likert scale from 1 (totally disagree) to 5 (totally agree), with higher scores indicating a stronger sense of self-threat in the situation. In Study 2, Cronbach's alpha was excellent ($\alpha = 0.89$), surpassing the reliability range (0.76–0.78) reported in the original validation study by Berjot et al. (2012).

Procedure

Study 1 was designed to explore the role played by ST in the emergence of depression. Based on the aforementioned theoretical elements which place threats to the self at the heart of depressive vulnerability, we propose a mediation model predicting that the relationship between DA and depressive symptoms will be mediated by ST. In Study 1, participants were recruited on social media (e.g., Facebook, LinkedIn) in November 2021 and were asked to complete a questionnaire administered online via the LimeSurvey platform. The participants,

drawn from the general population, indicated their age, biological sex, native language, employment sector, and completed the scales assessing dysfunctional attitudes (DA), self-threat (ST), and depressive symptoms.

Integrating both theoretical frameworks previously mentioned and results from Study 1, Study 2 was designed to provide evidence of the existence of a depressogenic prism through which individuals with DA appraise any situation. To that end, we randomly exposed students to one of two distinct scenarios (positive versus negative) and examined their degree of ST as a function of DA and scenario valence. After controlling the validity of our experimental manipulation, we investigated to what extent individuals feel threatened when facing a negative or a positive scenario depending on their level of DA. We argue that individuals with high level of DA will appraise both positive and negative situations as more threatening than individuals with low level of DA (H1). Reasoning that normal or adaptive levels of ST should still be expected in certain adverse circumstances, we fixed a "normal ST threshold" using ST levels of low-DA individuals in a negative situation as a baseline. In addition to first hypothesis, we go one step further and predict that even in a positive situation, individuals with high levels of DA will feel threatened above normal ST threshold (H2).

In Study 2, we built two distinct scenarios in which we invited students awaiting exam results to picture themselves in an imaginary yet potentially occurring situation. Scenarios were designed to bring into play the two dimensions identified by Beck (1967), namely performance evaluation and approval versus rejection by others: in each scenario, participants were not only asked to picture themselves succeeding or failing exams, they were also instructed to imagine the social ramifications of said results. Data were collected in a time-restricted period between the end of exams period (i.e., 14 January 2022) and communication of exam results (i.e., 2 February 2022). Thus, data were collected from 18 January to 28 January 2022 through online survey. All these elements (temporal, population, and scenario's content) contribute to make that scenarios as plausible as possible. Participants first indicated their age, gender, and university major, and were then presented with the DA scale, a randomly selected scenario (randomization was based on birthday date), ST scale, and a negative affect scale (always in this order).

Study 1 and Study 2 conformed with Helsinki Declaration (2008) and participation was completely voluntary. Participants were informed about (1) the general purpose of the study, (2) their rights to stop the survey at any time, and (3) the strict respect for confidentiality and anonymity. Prior to accessing the

questionnaire, all participants provided informed consent in accordance with the Helsinki Declaration.

Data Analysis

In Study 1, descriptive statistics, and Spearman's ρ correlation coefficient analysis were performed. Additionally, we evaluated the normality of all variables by calculating their Skewness and Kurtosis values. It is widely accepted that data is considered to follow a normal distribution when skewness falls within the range of ± 1 and kurtosis falls within ± 3. Then, using Process (Model 4; Hayes, 2018), we run a mediation analysis with a sample of 5,000 for bootstrapping and a 95% confidence interval, where the absence of zero reflects a significant effect.

In Study 2, Descriptive statistics, Spearman's ρ correlation coefficient analysis, analyses of variance (ANOVA), and independent-samples t-test were. In addition, we assessed the normality of all variables by calculating their Skewness and Kurtosis values. In both Study 1 and Study 2, all analyses were conducted using SPSS software (version 28.0).

Results

In Study 1, Means, medians, standard deviations, kurtosis, skewness, and correlation coefficients are reported in Table 1. As expected, the three variables were positively and significantly correlated. ST and DA exhibited a strong positive correlation, with ST showing a stronger association with depressive symptoms than DA. In addition, all three variables exhibited a normal distribution, as indicated by their Skewness and Kurtosis values.

With a mean depression score of 5.31 and a threshold of 8, the findings from Study 1 indicate that the sample experienced relatively low levels of depressive symptoms. Notably, only 76 out of 311 participants scored above the threshold, suggesting that most individuals in the sample did not exhibit significant depressive symptomatology.

Given the redundancies between some items of the DA scale (e.g., "I am nothing if a person I love doesn't love me") and the ST scale (e.g., "I am concerned that others will not approve of me"), as well as the strong correlation between these two scales, we calculated the Variance Inflation Factor (VIF, an index of multicollinearity). Results indicated that multicollinearity was not a concern (DA, Tolerance = 1, VIF = 1; ST, Tolerance = 1 VIF = 1; any value above 3 reflects collinearity). Therefore, even if they are highly correlated, DA and ST are two distinct constructs.

A mediation analysis was performed to assess the mediating role of ST in the relationship between DA and depressive symptoms (unstandardized coefficients are reported in Figure 1). The results confirmed the mediation model, with a significant indirect effect of DA on depressive symptoms through ST ($b = .2798$, $SE = .0451$, 95% CI [.191-.368], $p < .001$). The total effect of DA on depressive symptoms was significant ($b = .29506$, $SE = .0440$, 95% CI [.209-.381], $p < .001$). However, when controlling for ST, the direct effect of DA on depressive symptoms was no longer significant ($b = .0153$, $SE = .0590$, 95% CI [-.100-.131], $p = .796$). In other words, ST fully mediated the relationship between DA and depressive symptoms. The coefficients of each pathway are depicted in Figure 1.

It is noteworthy that consistent results were obtained in our preliminary study ($n = 115$) using the

Table 1. Means, medians, standard deviations, kurtosis, skewness, and intercorrelation between Study 1 variables ($N = 311$)

	HADS-d	DA	ST
Depressive symptoms (HADS-d)	–		
Dysfunctional attitudes (DA)	0.356***	–	
Self-threat (ST)	0.484***	0.718***	–
Mean	5.31	18.40	34.90
Median	5.00	18.00	35.00
Standard deviation	0.459	0.554	1.14
Kurtosis	0.136	-0.40	-0.737
Skewness	0.651	0.142	-0.085

Note. HADS-d = Depression subscale from the Hospital Anxiety and Depression Scale with scores ranging from 0 to 21, with a threshold score of 8; Items of DA are rated on a 4-point Likert scale, with score ranging from 9 to 36; Items of ST are rated on a 6-point Likert Scale with score ranging from 10 to 60; ***$p< .001$.
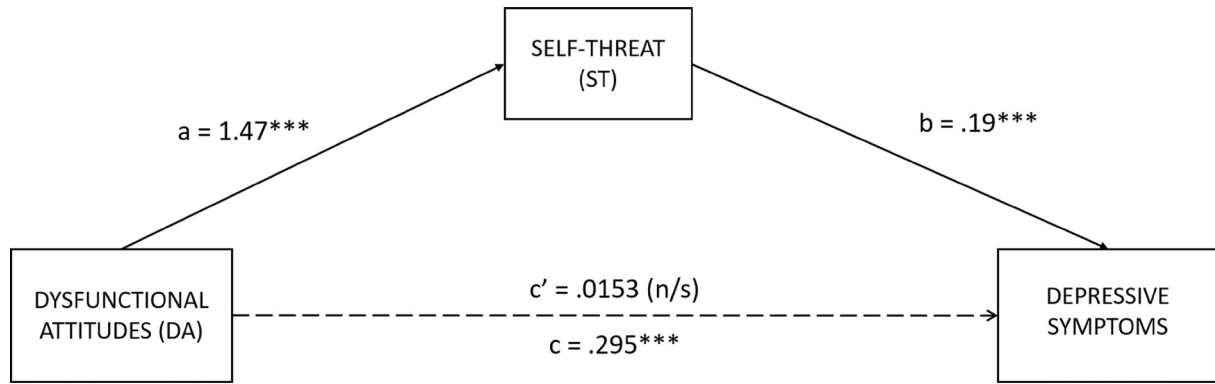
Figure 1. Mediation model of Study 1. Note. All coefficients are unstandardized regression weights; *a* is the effect of DA on ST; *b* is the effect of ST on depressive symptoms; *c'* is the direct effect of DA on depressive symptoms; *c* is indirect effect of DA on depressive symptoms through ST; ***$p < .001$.

same measure, except for the assessment of depressive symptomatology, which was conducted with the BDI-13 (Beck et al., 1974; French validation by Collet & Cottraux, 1986). Details of the mediation analysis results are presented in Appendix A.

In Study 2, Means, medians, standard deviations, kurtosis, skewness, and zero-order correlations are reported in Table 2. All our variables exhibited a normal distribution, as evidenced by the kurtosis and skewness values.

To control our experimental manipulation's validity, analyses of variance were conducted to examine the effects of the scenarios on negative emotions and self-threat state. The ANOVA revealed a significant effect of scenarios on both ST ($F(1) = 71.3$, $p < .001$; *Cohen's d* = 1.04) and on negative emotions ($F(1) = 80.4$, $p < .001$; *Cohen's d* = 1.11).

To explore how scenarios impacted ST as a function of DA, we split our sample into two groups of low and high DA, thus getting rid of the soft underbelly of our sample. DA groups are obtained by selecting the 68 individuals

who scored in the lower quartile of DA scale (Low$_{DAS}$ group; mean score$_{(DAS)}$ ≤ 1.78, comprising 42 females, 21 males and 5 non-binary), and the 74 individuals who scored in the upper quartile of DA scale (High$_{DAS}$ group; mean score$_{(DAS)}$ ≥ 2.56, comprising 53 females, 18 males and 3 non-binary). The two groups did not differ in terms of age [$t(140) = 0.922$, $p = 0.18$] or sex [$t(140) = 0.003$, $p = 0.50$]. A series of independent t-tests (Fig.2) were performed to assess differences in ST toward positive and negative scenarios. Confirming the depressogenic prism hypothesis (H1), compared with Low$_{DAS}$ group, High$_{DAS}$ individuals felt significantly more threatened in both negative [$t(69) = -5.59$, $p < .001$, *Cohen's d* = 1.35] and positive [$t(69) = -3.76$, $p < .001$, $d = 0.888$] scenarios (Figure 2). Consistent with hypothesis 2, High$_{DAS}$ ST levels in a positive scenario did not differ from the Low$_{DAS}$ ST levels in a negative scenario [$t(63) = -.182$, $p < .428$]. This lack of significant difference is indicative of an above-threshold ST in High$_{DAS}$ individuals even when exposed to a positive situation (Figure 2).

Table 2. Means, medians, standard deviations, kurtosis, skewness, and intercorrelation between Study 2 variables ($N = 263$)

	NA	DA	ST
Negative affect (NA)	–		
Dysfunctional Attitudes (DA)	0.34***	–	
Self-threat (ST)	0.74***	0.43***	–
Mean	27.80	19.70	17.40
Median	28.00	20.00	17.00
Standard deviation	10.80	5.07	6.72
Kurtosis	-1.04	-0.383	-1.00
Skewness	-0.076	0.225	0.081

Note. Items of NA are rated on a 5-point Likert scale with score ranging from 10 to 50; Items of DA are rated on a 4-point Likert scale, with score ranging from 9 to 36; Items of ST are rated on a 5-point Likert Scale with score ranging from 6 to 30; ***$p < .001$.
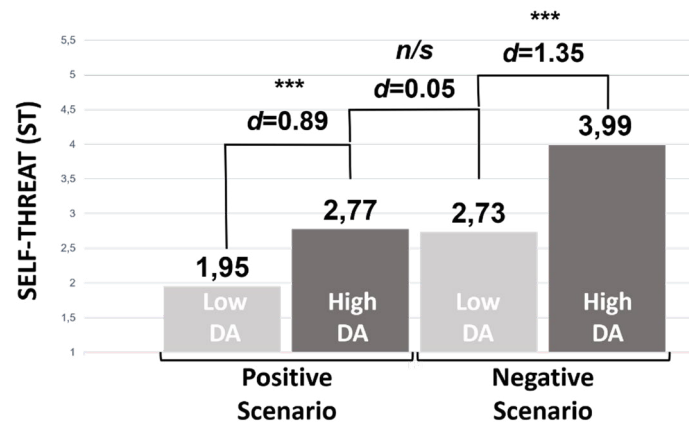
Figure 2. Differences in ST mean scores between individuals with low levels of DA (Low$_{DAS}$) and high levels of DA (High$_{DAS}$) according to valence scenario in Study 2. Note. DA = dysfunctional attitudes; ST scale comprises 5-point items (from 1 to 5); ***$p <$ .001; *n/s* = non-significant; d = Cohen's *d* (Cohen's *d* scores of 0.20, 0.50 and 0.80 are considered to reflect small, medium, and large effect sizes, respectively.

Discussion

The present paper sought to provide evidence for the centrality of ST in the emergence of depressive symptoms. To this end, we conducted two distinct studies in the general population to 1) investigate the mediating role of ST between DA and depressive symptoms, and 2) to test our depressogenic prism hypothesis which states that individuals with DA invariably appraise situations as threatening for their self, no matter the objective circumstances. The mediation analyses performed in Study 1 confirmed our first hypothesis: ST fully mediated the relationship between DA and depressive symptoms. It is worth noting that even though DA have commonly been designated as a main cause of depression in the scientific literature, correlations between ST and depressive symptoms (r = .484) were systematically greater than those between DA and depressive symptoms (r = .356). Study 2 involved direct emotional induction and was aimed at confirming our depressogenic prism hypothesis which proposes that individuals with high level of DA (High$_{DAS}$) appraise any situation as more threatening than individuals with low level of DA (Low$_{DAS}$). The existence of a depressogenic prism was confirmed as High$_{DAS}$ individuals felt significantly more threatened than Low$_{DAS}$ individuals in both scenarios. Interestingly, High$_{DAS}$ individuals felt as threatened in a positive situation as Low$_{DAS}$ individuals were in a negative situation. Results from Study 2 provide some original and significant data to the highly relevant question *"do some experiences always pose a threat to identity or does identity threat result from an individual's subjective appraisal of an experience?"* (Petriglieri,

2011, p.641). Here, we show that ST, a core mechanism in the emergence of depression, is directly dependent on individual subjectivity: in other words, the reality of individuals that fall prey to dysfunctional attitudes seems to invariably unfold under an ominous sky. It is therefore logical to assume that ST might be the actual end-product of DA and could therefore constitute the active agent responsible for the onset of depressive symptoms.

Certain limitations of this research merit mention. In our studies, we chose to adopt a transdiagnostic approach via the investigation of causal processes involved in the emergence of depressive symptoms in a general sample. We worked with the assumption that manifestations of depression vary along a continuum that comprises various degrees of severity, but that all those manifestations ultimately share a common structure and common root causes. Our main goal was not to pinpoint the exact mechanisms involved in the onset of the most extreme forms of acute depressive episodes, but rather to identify the most fundamental (and universally shared by human beings) causes of cognitive, affective, and behavioral manifestations of depression, even in their day-to-day and seemingly mundane form. However, a sample with low levels of depressive symptoms, as observed in Study 1, may limit the generalizability of the findings, rendering them less applicable to individuals with moderate to severe depression. This limitation also risks overlooking key mechanisms that are significant at higher levels of depression, potentially underestimating their role in the onset and progression of the condition. Therefore, further studies should explore the role of ST in the emergence of major depressive episodes among diagnosed patients.

Another limitation is that we integrated several theoretical accounts of the self and have thus considered identity and the self as two interchangeable concepts. However, one could argue that interchangeability was in fact conflation. Since self-related processes are obviously numerous and heterogeneous (Britton et al., 2021), future research should strive to identify the specific role of each self-related process involved in ST using distinct paradigms that isolate each of those processes. For instance, the well-researched body illusion (e.g., body swap illusion, rubber-hand illusion), in which participants experience an exchange of body with a dummy, could help targeting the effects of threats to the embodied self on depressive symptomatology, while other paradigms could target (without being exhaustive) the narrative self, the volitional self, or the social self.

Valuable for capturing insights into population characteristics, cross-sectional studies have limitations due to their static nature, making it challenging to establish causation or temporal relationships between variables. Although our hypotheses were confirmed by three separate studies (see Appendix A for the preliminary study), including one quasi-experimental study (i.e., Study 2), involving a large sample (*N*=689), it is crucial to complement our findings with longitudinal studies, providing a more nuanced understanding of the phenomena under investigation. This approach ensures a comprehensive exploration of the dynamic aspects of the studied variables over time.

Finally, this study relied solely on self-reported measures, which depend heavily on introspective abilities and social desirability. This limitation is particularly relevant here, as our study focused on subclinical depressive symptoms, which may be subtler and harder to detect. Getting back to the notion that depression is an evolved neuropsychobiological program mitigating social exclusion risks, it is plausible that its emergence involves both conscious perceptions and more implicit and elusive processes. Research has indeed shown that several evolutionary old processes contribute to various evaluative, emotional, and social responses, yet operate beyond the reach of conscious oversight (Elimari, 2023; Elimari & Lafargue, 2020, 2022, 2023a, 2023b, 2024). To capture these effects more comprehensively, objective measures (e.g., implicit, eye-tracking, electrodermal, neurophysiological) could offer additional insights. We might however contrast this limit by the yardstick of our findings: they reveal that depressive symptoms align more closely with individual perception than objective events, confirming that one cannot dispense with individual subjectivity when studying human behavior. Ultimately, even the most objective measures such as fMRI, diffusion tensor imaging, and even awake brain surgery techniques (e.g., direct electrical stimulations) rely on people's subjective reports to make sense of observed data (see Elimari & Lafargue, 2020, for an in depth discussion). In many ways, third person measures and phenomenology constitute two sides of the same coin.

Overall, we suggest that our findings offer valuable guidance for developing cognitive and emotional therapeutic strategies that could enhance preventive approaches to depression. Mindfulness-Based Interventions (MBIs) and Cognitive Behavioral Therapy (CBT), both with well-documented benefits (e.g., Gordillo Rodríguez et al., 2024), are recognized for their direct impact on dysfunctional attitudes (Britton et al., 2021), as well as for fostering cognitive decentering (Moore et al., 2022), promoting reappraisal (Garland et al., 2009), and enhancing mental clarity (Vago & Silbersweig, 2012). Both MBIs and CBT commonly employ techniques such as self-inquiry (or self-reflection) and discursive analysis as part of their therapeutic strategies (Dahl et al., 2015). Self-inquiry involves exploring one's thoughts, beliefs, and emotions to gain insight into their underlying patterns, which helps individuals recognize cognitive distortions and automatic responses that contribute to identity threat appraisal. Similarly, discursive analysis encourages the critical examination of one's inner dialogue and assumptions, enabling individuals to deconstruct rigid self-concepts and reduce the emotional impact of perceived identity threats. We suggest that these techniques could be particularly effective in addressing identity threat appraisal, as they may enhance self-awareness and promote greater cognitive flexibility in how individuals interpret and respond to their personal experiences, thereby supporting emotional regulation and facilitating cognitive restructuring.

Second-Generation Mindfulness-Based Interventions (SG-MBIs), which integrate spiritual and ethical components, have recently gained significant attention. As mentioned earlier, the Tibetan teacher Tsultrim Gyamtso suggests that all our suffering is birthed by an erroneous sense of self (i.e., autonomous, independent, permanent), which creates a feeling of threat and a constant preoccupation with protecting the self. Through SG-MBIs, one may discover that the sense of self as we conceive it (i.e., independent and permanent) is illusory and thus, this feeling of threat and this preoccupation to protect our self is ultimately meaningless (Varela et al., 2017). Alongside an understanding of Buddhist concepts such as impermanence, interdependence, and non-attachment, the careful analysis of beliefs and thoughts may dismantle one of the main foundations of depressive cognition, namely dysfunctional attitudes (DA). Moreo-

ver, not only do we perceive ourselves as permanent and independent, but we also attribute these same properties to our beliefs (e.g., a depressive individual may believe that statements like 'I'm a failure' or 'I'm unlovable' are universally true and apply in all contexts). Drawing on both traditional and scientific Buddhist literature, we propose that SG-MBIs could be particularly effective in addressing DA and ST to prevent the onset of depressive symptoms. Future research could examine the impact of SG-MBIs on these two risk factors, particularly within clinical populations.

Though we must remain cautious about the conclusions we draw from this paper, perhaps our results call for broader insights potentially relevant for western society as a whole. Indeed, with the knowledge that the feeling of ST inevitably calls for a response from individuals to protect their sense of self (Crocker & Park, 2004), we may suppose that perception of ST can lead people to claim more rights and recognition, or to vainly pursue more approval, material goods, money, or power (e.g., Dambrun & Ricard, 2011). Perhaps the answers our society provides (e.g., giving assistance or hardware solutions) ultimately makes the problem worse. If the research presented and cited in the present paper are accurate, depression is nothing more than the adaptive response to the danger that represents social disconnection. Though we do not dismiss the usefulness of large-scale public measures that secure people's material or financial safety, we believe that social measures tailored toward the maximization of social bonds might provide solutions that directly address the root cause of this elusive feeling of threat to the self, and thus depression as well as the non-adaptive behaviors that come along with it.

Conflicts of interest

The authors declare no conflicts of interest.

References

Allen, N. B., & Badcock, P. B. T. (2003). The Social Risk Hypothesis of Depressed Mood: Evolutionary, Psychosocial, and Neurobiological Perspectives. *Psychological Bulletin*, *129*(6), 887–913. https://doi.org/10.1037/0033-2909.129.6.887

Alloy, L. B., Abramson, L. Y., Hogan, M. E., Whitehouse, W. G., Rose, D. T., Robinson, M. S., Kim, R. S., & Lapkin, J. B. (2000). The Temple-Wisconsin Cognitive Vulnerability to Depression Project: Lifetime history of axis I psychopathology in individuals at high and low cognitive risk for depression. *Journal of Abnormal Psychology*, *109*(3), 403-418. http://dx.doi.org/10.1037/0021-843X.109.3.403

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, *117*, 497–529. https://doi.org/10.1037/0033-2909.117.3.497

Beck, A. T. (1967). *Depression*. New York.

Beck, A. T., & Bredemeier, K. (2016). A Unified Model of Depression: Integrating Clinical, Cognitive, Biological, and Evolutionary Perspectives. *Clinical Psychological Science*, *4*(4), 596–619. https://doi.org/10.1177/2167702616628523

Beck, A. T., Rial, W. Y., & Rickels, K. (1974). Short form of depression inventory: Cross-validation. *Psychological Reports*, *34*(3_suppl), 1184–1186. https://doi.org/10.1177/003329417403403s01

Beevers, C. G., Strong, D. R., Meyer, B., Pilkonis, P. A., & Miller, I. W. (2007). Efficiently assessing negative cognition in depression: An item response theory analysis of the Dysfunctional Attitude Scale. *Psychological Assessment*, *19*(2), 199. https://doi.org/10.1037/1040-3590.19.2.199

Berjot, S., & Bourguignon, D. (2023). *Le Soi et l'identité en psychologie sociale: Fondements, concepts et applications*. Dunod. https://doi.org/10.3917/dunod.berjo.2023.01

Berjot, S., & Girault-Lidvan, N. (2009). Validation d'une version française de l'échelle d'évaluation cognitive primaire de Brewer et Skinner. *Canadian Journal of Behavioural Science / Revue Canadienne Des Sciences Du Comportement*, *41*(4), 252–259. https://doi.org/10.1037/a0014842

Berjot, S., Girault-Lidvan, N., & Gillet, N. (2012). Appraising Stigmatization and Discrimination: Construction and Validation of a Questionnaire Assessing Threat and Challenge Appraisals to Personal and Social Identity. *Identity*, *12*(3), 191–216. https://doi.org/10.1080/15283488.2012.691254

Bjelland, I., Dahl, A. A., Haug, T. T., & Neckelmann, D. (2002). The validity of the Hospital Anxiety and Depression Scale: An updated literature review. *Journal of Psychosomatic Research*, *52*(2), 69–77. https://doi.org/10.1016/S0022-3999(01)00296-3

Bocéréan, C., & Dupret, E. (2014). A validation study of the Hospital Anxiety and Depression Scale (HADS) in a large sample of French employees. *BMC Psychiatry*, *14*(1), 354. https://doi.org/10.1186/s12888-014-0354-0

Bouvard, M., Cottraux, J., & Charles, S. (1994). Etude de validation sur une population française de l'échelle d'attitudes dysfonctionnelles de Weissman et Beck (DAS-forme A). *Journal de Thérapie Comportementale et Cognitive*, *4*, 127–135. http://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=3328018

Britton, W. B., Desbordes, G., Acabchuk, R., Peters, S., Lindahl, J. R., Canby, N. K., Vago, D. R., Dumais, T., Lipsky, J., & Kimmel, H. (2021). From self-esteem to selflessness: An evidence (gap) map of self-related processes as mechanisms of mindfulness-based interventions. *Frontiers in Psychology*, *12*, 730972. https://doi.org/10.3389/fpsyg.2021.730972

Bstan-vdzin-rgya-mtsho (1997). *The four noble truths: Fundamentals of the Buddhist teachings*. Thorsons.

Camuñas, N., Mavrou, I., y Miguel-Tobal, J. J. (2019). Ansiedad y tristeza-depresión: Una aproximación desde la teoría de la indefensión-desesperanza. *Revista de Psicopatología y Psicología Clínica*, *24*(1), 19-28. https://doi.org/10.5944/rppc.23003

Cohen, G. L., Purdie-Vaughns, V., & García, J. (2012). An identity threat perspective on intervention. In *Stereotype threat: Theory, process, and application* (pp. 280–296). Oxford University Press. 10.1093/acprof:oso/9780199732449.003.0018

Collet, L., & Cottraux, J. (1986). The shortened Beck depression inventory (13 items). Study of the concurrent validity with the Hamilton scale and Widlöcher's retardation scale. *L'encephale*, *12*(2), 77–79. http://pascal-francis.inist.fr/vibad/index.php?action=getRecordDetail&idt=8027038

Crocker, J., & Park, L. E. (2004). The costly pursuit of self-esteem. *Psychological Bulletin*, *130*(3), 392. https://doi.org/10.1037/0033-2909.130.3.392

Dahl, C. J., Lutz, A., & Davidson, R. J. (2015). Reconstructing and deconstructing the self: Cognitive mechanisms in meditation practice. *Trends in Cognitive Sciences*, *19*(9), 515–523. https://doi.org/10.1016/j.tics.2015.07.001

Dambrun, M., & Ricard, M. (2011). Self-centeredness and selflessness: A theory of self-based psychological functioning and its consequences for happiness. *Review of General Psychology*, *15*(2), 138–157. https://doi.org/10.1037/a0023059

Dickerson, S. S., Gruenewald, T. L., & Kemeny, M. E. (2009). Psychobiological Responses to Social Self Threat: Functional or Detrimental? *Self and Identity*, *8*(2–3), 270–285. https://doi.org/10.1080/15298860802505186

Disner, S. G., Beevers, C. G., Haigh, E. A., & Beck, A. T. (2011). Neural mechanisms of the cognitive model of depression. *Nature Reviews Neuroscience*, *12*(8), 467–477. https://doi.org/10.1038/nrn3027

Elimari, N. (2023). *Approches duelle et évolutive de la neurocognition: Nature et corrélats neuraux des processus de régulation entre instincts et cognition de haut-niveau* (Doctoral dissertation). Reims.

Elimari, N., & Lafargue, G. (2020). Network neuroscience and the adapted mind: rethinking the role of network theories in evolutionary psychology. *Frontiers in Psychology*, *11*, 545632. https://doi.org/10.3389/fpsyg.2020.545632

Elimari, N., & Lafargue, G. (2022). Influence du système immunitaire comportemental sur la xénophobie et l'altruisme en temps de pandémie: Le cas de la crise COVID-19. *Les épidémies au prisme des SHS*, 75-82. https://doi.org/10.17184/eac.5992

Elimari, N., & Lafargue, G. (2023a). Neural correlates of performance monitoring vary as a function of competition between automatic and controlled processes: An ERP study. *Consciousness and Cognition*, *110*, 103505. https://doi.org/10.1016/j.concog.2023.103505

Elimari, N., & Lafargue, G. (2023b). Qu'est-ce que la psychologie évolutive? Cognition et comportements à la lumière de l'évolution. *L'Année psychologique*, *123*(1), 173-214. https://doi.org/10.3917/anpsy1.231.0173

Elimari, N., & Lafargue, G. (2024). Two social minds in one brain? error-related negativity provides evidence for parallel processing pathways during social evaluation. *Cognition and Emotion*, *38*(1), 90-102. https://doi.org/10.1080/02699931.2023.2270200

Espejo, E. P., Hammen, C., & Brennan, P. A. (2012). Elevated Appraisals of the Negative Impact of Naturally Occurring Life Events: A Risk Factor for Depressive and Anxiety Disorders. *Journal of Abnormal Child Psychology*, *40*(2), 303–315. https://doi.org/10.1007/s10802-011-9552-0

Gadassi, R., & Rafaeli, E. (2015). Interpersonal perception as a mediator of the depression–interpersonal difficulties link: A review. *Personality and Individual Differences*, *87*, 1–7. https://doi.org/10.1016/j.paid.2015.07.023

Garland, E., Gaylord, S., & Park, J. (2009). The Role of Mindfulness in Positive Reappraisal. *EXPLORE*, *5*(1), 37–44. https://doi.org/10.1016/j.explore.2008.10.001

Gaudreau, P., Sanchez, X., & Blondin, J.-P. (2006). Positive and negative affective states in a performance-related setting: Testing the factorial structure of the PANAS across two samples of French-Canadian participants. *European Journal of Psychological Assessment*, *22*(4), 240–249. https://doi.org/10.1027/1015-5759.22.4.240

Gordillo Rodríguez, R., González-Peña, P., del Barrio, V., y Ramos, R. (2024). Impact of COVID-19 pandemic on videoconferencing psychotherapies for the treatment of depression in Spain. *Revista de Psicopatología y Psicología Clínica*, *29*(1), 113–120. https://doi.org/10.5944/rppc.37998

Gunthert, K. C., Cohen, L. H., & Armeli, S. (1999). The role of neuroticism in daily stress and coping. *Journal of Personality and Social Psychology*, *77*(5), 1087. https://doi.org/10.1037/0022-3514.77.5.1087

Hayes, A. F. (2018). Partial, conditional, and moderated moderated mediation: Quantification, inference, and interpretation. *Communication Monographs*, *85*(1), 4–40. https://doi.org/10.1080/03637751.2017.1352100

Kopala-Sibley, D. C., & Zuroff, D. C. (2020). The self and depression: Four psychological theories and their potential neural correlates. *Journal of Personality*, *88*(1), 14–30. 10.1111/jopy.12456

Kuiper, N. A., Olinger, L. J., & Martin, R. A. (1988). Dysfunctional attitudes, stress, and negative emotions. *Cognitive Therapy and Research*, *12*(6), 533–547. https://doi.org/10.1007/BF01205008

Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford University Press. https://doi.org/10.2307/2075902

Moore, M. T., Lau, M. A., Haigh, E. A., Willett, B. R., Bosma, C. M., & Fresco, D. M. (2022). Association between decentering and reductions in relapse/recurrence in mindfulness-based cognitive therapy for depression in adults: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *90*(2), 137. 10.1037/ccp0000718

Petriglieri, J. L. (2011). Under Threat: Responses to and the Consequences of Threats to Individuals' Identities. *Academy of Management Review*, *36*(4), 641–662. https://doi.org/10.5465/amr.2009.0087

Pietromonaco, P. R., & Markus, H. (1985). The nature of negative thoughts in depression. *Journal of Personality and Social Psychology*, *48*, 799–807. https://doi.org/10.1037/0022-3514.48.3.799

Rainville, P., Duncan, G. H., Price, D. D., Carrier, B., & Bushnell, M. C. (1997). Pain affect encoded in human anterior cingulate but not somatosensory cortex. *Science*, *277*(5328), 968–971. 10.1126/science.277.5328.968

Ruiz-Alonso, E., Orue, I., y Calvete, E. (2021). Relaciones

bidireccionales longitudinales entre victimización, estilos inferenciales de desesperanza y síntomas de depresión en adolescentes: Un modelo transaccional. *Revista de Psicopatología y Psicología Clínica*, *26*(2), 121–130. https://doi.org/10.5944/rppc.28778

Ryan, R. M., & Deci, E. L. (2002). Overview of self-determination theory: An organismic dialectical perspective. *Handbook of Self-Determination Research*, *2*, 3–33.

Scher, C. D., Ingram, R. E., & Segal, Z. V. (2005). Cognitive reactivity and vulnerability: Empirical evaluation of construct activation and cognitive diatheses in unipolar depression. *Clinical Psychology Review*, *25*(4), 487–510. 10.1016/j.cpr.2005.01.005

Şenormancı, Ö., Yılmaz, A. E., Saraçlı, Ö., Atasoy, N., Şenormancı, G., & Atik, L. (2014). The mediator role of ruminative thinking style in the relationship between dysfunctional attitudes and depression. *Comprehensive Psychiatry*, *55*(7), 1556–1560. https://doi.org/10.1016/j.comppsych.2014.05.017

Sharma, S., & Sharma, M. (2010). Self, social identity and psychological well-being. *Psychological Studies*, *55*(2), 118–136. https://doi.org/10.1007/s12646-010-0011-8

Skinner, N., & Brewer, N. (2002). The dynamics of threat and challenge appraisals prior to stressful achievement events. *Journal of Personality and Social Psychology*, *83*(3), 678. https://doi.org/10.1037/0022-3514.83.3.678

Vago, D. R., & Silbersweig, D. A. (2012). Self-awareness, self-regulation, and self-transcendence (S-ART): A framework for understanding the neurobiological mechanisms of mindfulness. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00296

Varela, F. J., Thompson, E., & Rosch, E. (2017). *The embodied mind, revised edition: Cognitive science and human experience*. MIT press.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063. 10.1037/0022-3514.54.6.1063

Weissman, A. N., & Beck, A. T. (1978). *Development and validation of the Dysfunctional Attitude Scale: A preliminary investigation*. Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canadá, March 27-21, 1978)

Wetherall, K., Robb, K. A., & O'Connor, R. C. (2019). Social rank theory of depression: A systematic review of self-perceptions of social rank and their relationship with depressive symptoms and suicide risk. *Journal of Affective Disorders*, *246*, 300–319. https://doi.org/10.1016/j.jad.2018.12.045

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, *67*(6), 361–370. https://doi.org/10.1111/j.1600-0447.1983.tb09716.x

Appendix A

A preliminary study was conducted to explore the role played by ST in the emergence of depression. A total of 115 workers (women 67%, *M* age = 36 years, *SD* = 11.56) from various organizations completed the online survey which included the Dysfunctional Attitudes Scale (DAS-9; Beevers et al., 2007), the Threat Appraisal Subscale of the Cognitive Appraisal Scale (Skinner & Brewer, 2002), and the 13-items Beck Depression Inventory (BDI-13; Beck et al., 1974). Using Process (Model 4; Hayes, 2018), a mediation analysis was run with a sample of 5,000 for bootstrapping and a 95% confidence interval. Results confirmed the mediation model (see figure a), with a significant indirect effect of DA on depressive symptoms through ST ($b = .3821$, $SE = .0887$, 95% *CI* [.2115-.5646], $p < .001$). The total effect of DA on depressive symptoms was significant ($b = .5362$, $SE = .0835$, 95% *CI* [.3707-.7017], $p < .001$). However, when controlling for ST, the direct effect of DA on depressive symptoms was no longer significant ($b = .1541$ $SE = .0978$, 95% *CI* [-.0397-.3480], $p = .118$). In other words, ST fully mediated the relationship between DA and depressive symptoms.
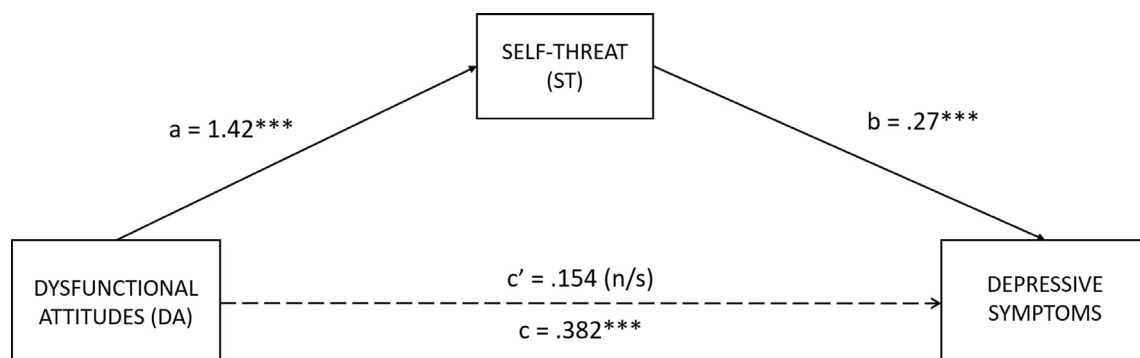

Figure a. Mediation model of our preliminary study. Note. All coefficients are unstandardized regression weights; *a* is the effect of DA on ST; *b* is the effect of ST on depressive symptoms; *c'* is the direct effect of DA on depressive symptoms; *c* is indirect effect of DA on depressive symptoms through ST; ***$p < .001$; *n/s* = non-significant with $p = .118$.