TOPIC - Temi di Psicologia dell'Ordine degli Psicologi della Campania
V. 5 N. 2 (2026)
https://doi.org/10.53240/2026TOPIC5N2.111

SOCRATE vs. pensieri autocritici: risultati preliminari su un'Applicazione per un Intervento basato sul Dialogo Socratico

SOCRATE vs. critical thoughts: Preliminary results on an Application for Socratic Dialogue Intervention

Vanda Lamberti[1], Francesco Caputo[1], Elena Piscopo[1], Guido Almaviva[1], Alessio Lustro[1], Alessandro Couyoumdjian[2], Micaela Di Consiglio[1,2], Vittoria Zaccari[1,2,3,4], Amelia Gangemi[5], Francesco Mancini[4], Angelo Maria Saliani[4], Francesca D'Olimpio[1]

[1] *Dipartimento di Psicologia, Università degli studi della Campania "Luigi Vanvitelli", Caserta, Italia*

[2] *Dipartimento di Psicologia, Sapienza Università di Roma, Roma, Italia*

[3] *Dipartimento di Scienze Umane, Università Guglielmo Marconi, Roma, Italia*

[4] *Associazione di Psicologia Cognitiva APC e Scuola di Psicoterapia Cognitiva SPC, Roma, Italia*

[5] *Dipartimento di Scienze Cognitive, Università di Messina, Messina, Italia*

Abstract

L' autocritica è un processo transdiagnostico associato a un aumento del disagio psicologico e a esiti terapeutici sfavorevoli, in particolare nella popolazione con tratti ossessivo-compulsivi caratterizzati da rigide autovalutazioni morali e dialoghi interiori di tipo punitivo. Il seguente studio pilota ha esaminato la fattibilità e gli effetti preliminari di SOCRATE in un campione non clinico. SOCRATE è un'applicazione finalizzata a fornire un intervento digitale strutturato basato sulla tecnica del dialogo socratico, rivolto specificamente all'architettura dialogica e inferenziale dei processi autocritici. Sedici giovani adulti non clinici (N=16) hanno completato una valutazione attraverso misure relative alla sintomatologia del Disturbo Ossessivo-Compulsivo (DOC) e dell'ansia sociale prima e dopo la somministrazione di SOCRATE, utilizzando l'applicazione quotidianamente per quindici giorni. I risultati evidenziano una riduzione significativa della sintomatologia ossessivo-compulsiva tra i partecipanti non clinici che presentavano tratti DOC nel pre-test. Non sono emersi cambiamenti significativi per gli altri domini sintomatici. I risultati preliminari suggeriscono che il dialogo socratico digitale e strutturato possa rappresentare un nuovo intervento a bassa intensità sui processi autocritici associati alla sintomatologia legata al DOC. Inoltre, nel panorama della salute mentale digitale, questo lavoro contribuisce ad accrescere le evidenze scientifiche circa gli interventi digitali strutturati come strumenti in grado di potenziare gli effetti della psicoterapia, in particolare quelli mirati alla modulazione e alla riduzione dell'autocritica. Ulteriori studi saranno necessari per determinare e chiarire come tali strumenti dialogici a bassa intensità possano essere integrati in modo ottimale all'interno di più ampi modelli di intervento psicologico: in particolare, valutare l'efficacia dell'uso combinato di SOCRATE con trattamenti psicoterapeutici e stabilire le condizioni sotto le quali tale integrazione possa produrre il maggior beneficio clinico.

V. 5 N. 2 (2026)
https://doi.org/10.53240/2026TOPIC5N2.111

Parole Chiave
Autocritica, Tratti Ossessivo-Compulsivi, Dialogo Socratico, Interventi Digitali, Studio Pilota, Processi Transdiagnostici.

Abstract
Self-Criticism (SC) is a transdiagnostic process associated with increased psychological distress and poor therapeutic outcomes, particularly among individuals with obsessive-compulsive traits characterized by rigid moral self-evaluations and punitive inner dialogues. This pilot study examined the feasibility and preliminary effects of SOCRATE in a non-clinical sample. SOCRATE consists of an App aimed at providing a structured digital intervention based on the Socratic Dialogue technique, which is addressed specifically to the dialogic and inferential architecture of self-critical processes. Sixteen non-clinical young adults (N=16) completed an assessment of measures, before and after the intervention referred to Obsessive Compulsive Disorder and Social Anxiety symptoms and then they used the application daily for fifteen days. Results revealed a significant reduction in obsessive-compulsive symptomatology among non-clinical participants who presented OCD traits in the pre-test measure. No significant changes emerged for other symptomatic domains. These preliminary findings suggest that structured digital Socratic Dialogue may represent a promising low-intensity intervention for self-critical processes associated with OCD-related symptomatology. Furthermore, within the scene of digital mental health, this work contributes to the growing body of research on structured digital interventions as a tool capable of enhancing the effects of psychotherapy, particularly those aimed at the modulation and reduction of SC. Further research is needed to determine and clarify how such low-intensity dialogical tools can be optimally integrated into broader models of psychological intervention. In particular, additional studies are required to evaluate the effectiveness of combining SOCRATE usage integrated into psychotherapy and to establish the conditions under which this integration may yield the greatest clinical benefit.

Keywords
Self-Criticism, Obsessive-compulsive traits, Socratic Dialogue, Digital intervention, Pilot study, Transdiagnostic processes.

Introduction

Self-Criticism (SC) has been defined as an integrated system of beliefs, emotions, and attitudes that people might have toward themselves in response to failures or setbacks (Gilbert, 2005). It operates as an internal process in which individuals judge and attack themselves through a hostile inner dialogue, often involving negative, coercive self-talk and punitive internal voices (Longe et al., 2010). This suggests that SC is organized not only in terms of cognitive content but also in a dialogical form.

Consistent with Beck's cognitive model (1967, 1976, 1979), the self-critical process is directly sustained by cognitive distortions, such as personalization, overgeneralization, disqualifying the positive, and selective abstraction. These biases amplify perceived failures and setbacks while minimizing personal qualities and successes. Such distortions activate

Autore responsabile per la corrispondenza: Vanda Lamberti, Dipartimento di Psicologia, Università degli studi della Campania "Luigi Vanvitelli", Caserta, Italia; e-mail: vanda.lamberti@unicampania.it

https://doi.org/10.53240/2026TOPIC5N2.111

and maintain a persistent stream of self-critical thoughts, where the self is interpreted as inadequate, guilty, or flawed through harsh and inflexible judgments. Within this framework, SC is experienced not merely as an emotion but as a recurring pattern of biased self-interpretations that become a stable component of a dysfunctional cognitive style.

Converging evidence indicates that elevated SC is associated with greater symptom severity (e.g., more severe depressive and anxiety symptoms, pervasive negative self-evaluation, and increased ruminative thinking), higher comorbidity, and poorer functional outcomes (Williams et al., 2022; Miceli & Castelfranchi, 2018; Golestaneh et al., 2017). Given the transdiagnostic nature of SC, targeting it during psychotherapy sessions appears clinically warranted. However, elevated SC has also been consistently associated with poorer psychotherapy outcomes, including reduced treatment responsiveness and an increased risk of dropout (Löw et al., 2020). This has stimulated growing interest in identifying interventions capable of effectively addressing self-critical processes.

Different aspects of SC are likely to be associated with distinct, or even relatively "pure," psychological profiles. SC grounded in moral criteria may be particularly involved in disorders in which guilt plays a central role, such as obsessive–compulsive disorder (OCD) and depression. Conversely, SC focused on one's personal qualities may be more closely linked to the emotion of shame, as observed in conditions such as eating disorders or social anxiety (Zaccari et al., 2024).

A new model proposed by Zaccari & Mancini (2026) further illustrates how different forms of inner dialogue may underlie distinct emotional responses. Within obsessive–compulsive symptomatology, the fear of deontological guilt (Mancini & Gangemi, 2021; Mancini et al., 2021) is typically expressed through a morally oriented critical inner dialogue. This dialogue is characterized by the anticipation of external criticism and manifests as a negative self-evaluation in which the individual perceives themselves as bad, wrong, or immoral (Mancini, 2019).

In social anxiety disorder and eating disorders, SC appears to be predominantly polarized towards a domain that does not concern morality (non-moral). Here, criticism is self-focused and revolves around themes of incompetence, inadequacy, unworthiness, or self-directed hatred and disgust. These evaluations are grounded in non-moral criteria—such as lovability, competence, intelligence, physical appearance, performance, and social inclusion—and tend to elicit emotions such as anxiety, shame, and embarrassment (Fairburn et al., 2003; Werner et al., 2011). In this form, SC stems from comparisons based on non-moral reference standards, including perceived intelligence, abilities, physical prowess, or performative capacity.

Within this framework, dialogical approaches conceptualize SC as an internalized relational dynamic rather than a static self-evaluative trait, allowing individuals to actively engage with and regulate critical inner voices through structured interpersonal-like exchanges.

SC is a transdiagnostic factor, not exclusively related to a clinical phenomenon but as a widespread process relevant to the general population, opening up the possibility of developing scalable interventions. For instance, Cerea et al. (2020) assessing the efficacy of short, game-like, daily cognitive interventions delivered via mobile application, evidenced a significant decrease in Obsessive Compulsive symptoms, social anxiety symptoms and a greater increase in

https://doi.org/10.53240/2026TOPIC5N2.111

self-esteem in participants using the application for 15 days.

Considering this literature, the present study examines the feasibility and preliminary effects of SOCRATE, a structured digital application designed to intervene in self-critical processes through a guided Socratic Dialogue (SD) protocol.

In a recent meta-analysis by Löw et al. (2020) the authors found that higher levels of SC are related to poorer outcome. Therefore, several authors have attempted to develop techniques for treating and directly targeting SC. Cognitive–behavioral therapies (CBT) and Compassion-Focused Therapy (CFT; Gilbert, 2014) have demonstrated efficacy in reducing SC in the short term, while also highlighting the need to refine existing protocols to more systematically address self-critical processes (Löw et al., 2020). For instance, in the CFT (Gilbert, 2014) patients are asked to become aware of the content and tone of their "internal dialogue" and to consciously make it more compassionate, encouraging, and nonjudgmental. Therefore, results on their efficacy are inconsistent and most studies investigating the association between SC and therapy outcomes have not discriminated between different aspects of SC. Low et al. (2020) pointed out that cognitive-behavioral therapies may be suitable for the short-term treatment of self-critical individuals, suggesting an improvement of existing methods with ingredients specifically targeting SC.

Overall, brief interventions appear effective in producing short-term reductions in SC, whereas evidence for long-term efficacy remains limited, prompting increasing interest in alternative or complementary approaches (Pekin et al., 2025).

Within CBT, SD represents a core mechanism for addressing maladaptive self-evaluations (Overholser, 2018). Conceptualized as a set of strategic questions integrated with guided discovery, SD is not merely a questioning technique but a collaborative process of change (Overholser, 2010; Padesky, 1993). Given that SC is often maintained through rigid and accusatory self-evaluative dialogues, SD seems particularly well suited to engaging with and exploring the underlying structure of such processes. SD facilitates a methodical examination and inductive restructuring of maladaptive appraisals through guided questioning covering four specific domains: logical-empirical (Is that true? / What evidence do you have?), pragmatic (Does it work?), economic (Is it worth it?), moral (Is it fair?). It promotes reflective awareness and the exploration of automatic thoughts by helping clients access and critically examine their own underlying assumptions and neglected experiences (Padesky, 1993), ultimately guiding the client toward generating more functional, evidence-based interpretations (Kazantzis et al., 2014).

Although the specific contribution of SD as an isolated technique remains difficult to disentangle empirically (Clark & Egan, 2015), converging clinical and theoretical evidence suggests that SD is especially relevant in conditions characterized by heightened SC and guilt (Saliani et al., 2024), given its capacity to engage clients in a structured exploration of rigid self-evaluative cognitions (Overholser, 2013).

The procedural organization of SD, based on rule-governed inferential steps and sequenced questioning, renders its core mechanisms transferable to formalization and systematic implementation, including digital applications. Recent evidence suggests that CBT-based digital interventions can successfully transfer key mechanisms of cognitive change to such settings, yielding meaningful effects even outside

https://doi.org/10.53240/2026TOPIC5N2.111

traditional therapeutic environments (Im & Woo, 2025).

Collectively, these findings suggest that structured dialogical mechanisms of CBT can be coherently translated into digital formats, yet they also underscore that the architecture of implementation is not a neutral design choice. Generative systems introduce specific concerns related to controllability, content safety, and the preservation of coherent clinical logic (Blease & Torous, 2023; Iftikhar et al., 2025). Such concerns are particularly salient when intervening in emotionally sensitive and self-evaluative processes such as SC.

Accordingly, SOCRATE was developed as a rule-based digital application that operationalizes selected core components of SD within a predefined inferential sequence. Rather than simulating an open-ended conversational exchange, the application guides users through a structured progression of Socratic prompts designed to facilitate the examination and restructuring of self-critical internal dialogues across both moral and non-moral domains. This constrained dialogical structure is particularly relevant when targeting SC, a process prone to escalation and perseverative self-attack, as it allows the intervention to remain focused, contained, and clinically coherent.

Within the broader landscape of digital mental health, this work contributes to the growing body of research on structured digital mental health interventions. Further research will be required to determine the durability of observed effects and to clarify how such low-intensity dialogical tools may be optimally integrated within broader models of psychological intervention.

To address the complexity of self-critical processes, SOCRATE integrates two distinct but complementary functional modules: a structured Dialogue with the Inner Critic and an Approach-Avoidance Task (AAT).

Dialogue with the Inner Critic operationalizes the inferential sequence of Socratic questioning. Users engage with a "Critic avatar" and respond using predefined Socratic prompts. While, Approach-Avoidance Task (AAT), is a module that follows the Dialogue with the Inner Critic and in which participants perform a gamified AAT. This behavioral paradigm is utilized to assess and potentially modify implicit behavioral tendencies (approach vs. avoidance) toward emotional stimuli (adjectives) related to the self-critical domains previously explored in the module Dialogue with the Inner Critic.

While SOCRATE is designed as a comprehensive transdiagnostic tool addressing both moral SC (prevalent in OCD and depression) and non-moral SC (typical of social anxiety and eating disorders), the primary objective of this pilot paper is to present preliminary evidence specifically focused on obsessive-compulsive symptomatology and social anxiety symptomatology. Although data regarding app usage were collected, to ensure the feasibility of the overall app architecture, this initial paper focuses on the application's impact on OCD traits and social anxiety symptoms. By isolating these preliminary findings, we aim to validate the core mechanism of the digital Socratic intervention before expanding the analysis to the broader transdiagnostic clinical spectrum.

Methods

Participants

Participants were 26 non-clinical young adults (between 18 and 35 years) recruited from university students enrolled in master's degree courses in the Department of Psychology at the University of Campania "Luigi Vanvitelli". Their participation was voluntary. The final sample

https://doi.org/10.53240/2026TOPIC5N2.111

comprised 16 participants (12 female, 4 males; *M*_age = 24.10, *SD* = 2.35), while 10 participants were excluded as they did not complete the protocol properly (e.g. they missed one or more daily app use). Participants' had bachelor's degree as educational level. All participants provided informed consent prior to participation, and the study was conducted in accordance with the Declaration of Helsinki.

Measures

Obsessive-compulsive symptomatology was assessed using the Obsessive–Compulsive Inventory–Revised (OCI-R; Foa et al., 2002; Italian validation by Marchetti et al., 2010), consisting of 18 items on a 5-point Likert scale ranging from 0 ("not at all") to 4 ("extremely"). The questionnaire evaluates frequency, characteristics and severity of OC symptomatology across six dimensions: washing, checking, ordering, hoarding, mental neutralizing, and obsessing. Higher scores indicate greater symptom severity. In the present study, the total score was used for analyses. Clinically significant obsessive-compulsive symptomatology was defined according to standardized criteria (Jensen, 2003), with a cut-off score of ≥ 21 on the total OCI-R, a threshold also considered indicative of a likely OCD diagnosis (Foa et al., 2002).

In the present sample, internal consistency for the total score was good at both time points (Pre α = .79; Post α = .81).

Social anxiety was measured using the Social Interaction Anxiety Scale (SIAS; Mattick & Clarke, 1998; Italian validation by Sica et al., 2007), a 20-item self-report questionnaire in which the participant rates how well they feel on a 5-point Likert scale, ranging from 0 ("not at all") to 5 ("extremely"). It specifically investigates the fear of interacting in social situations and measures the emotional aspects of anxiety. The total score, reflecting discomfort and anxiety in social interactions, was used in analyses. Cronbach's alpha in the Italian validation study was .86 for the total score while in this study internal consistency was also good (Pre α = .87; Post α = .83).

Manipulation Check

We assessed participants' state levels of emotions, self-criticism, self-compassion and self-esteem, through the entire experimental phase to determine whether the induction of self-criticism and the Socratic questioning were effective.

The choice to include self-criticism, self-compassion, and self-esteem is grounded in a well-established theoretical framework regarding self-regulation processes and self-functioning. Specifically, self-compassion was conceptualized by Neff (2003) as a positive and non-judgmental attitude toward oneself, articulated through the components of self-kindness, common humanity, and mindfulness, and is consistently associated with improved psychological well-being. Conversely, self-criticism has been widely identified as a psychopathological vulnerability factor, closely linked to emotions of shame, anxiety, and depression (Gilbert, 2000; Gilbert & Irons, 2005). It can be viewed as a dysfunctional mode of relating to the self, standing in direct opposition to the adaptive processes represented by self-compassion.

Finally, self-esteem represents a global evaluation of one's personal worth, traditionally defined as a positive or negative attitude toward oneself (Rosenberg, 1965). While this construct is a central dimension of self-functioning, it remains more closely tied to evaluative and comparative processes, making it more sensitive to external contingencies (Leary et al., 1995). Literature further highlights that self-esteem and self-compassion are distinct yet complementary:

https://doi.org/10.53240/2026TOPIC5N2.111

while the former reflects a global self-evaluation, the latter represents a more stable way of relating to oneself that is less dependent on social comparison (Neff & Vonk, 2009).

In line with this literature, the joint inclusion of these three domains allows for a multifaceted assessment of self-regulation processes by integrating dimensions of vulnerability (self-criticism), adaptation (self-compassion), and global evaluation (self-esteem). To empirically evaluate these constructs and determine the effectiveness of the self-criticism induction and Socratic questioning, we assessed participants' state levels of these variables throughout the entire experimental phase. For this purpose, we developed an ad hoc questionnaire consisting of 10 items assessing emotions, two items assessing self-compassion and self-criticism and the Single Item Self Esteem (SISE; Robins et al., 2001) "I have high self-esteem", for self-esteem measurement. Participants were asked to rate their emotional level on a 7 point Likert scale, from 0 ("not at all") to 6 ("extremely") and the ten emotions assessed in the following order: shame, joy, fear, disgust, sadness, guilt toward another person, anger, pity, moral guilt, boredom. Only at the first administration there was a definition of each emotion and a typically associated facial expression provided. To assess self-criticism and self-compassion, two items presented as the emotions described above were presented: "How self-critical are you feeling right now?" and "How self-compassionate are you feeling right now?". The items were also used as part of the manipulation check procedure in a previous study by Nagy et al. (2021). Finally global state self-esteem was assessed using the SISE.

App description

The application was developed by Immensive SRL for both Android and iOS devices to ensure cross-platform accessibility and minimize participant exclusion. Upon registration, participants provided basic demographic information (name, age, gender, the last three digits of their phone number) which was used to generate an anonymous alphanumeric code linking app usage data with survey responses while preserving anonymity.

Participants then created a personalized avatar by selecting from six base figures (three female and three male), each available in three skin tone options. Avatar customization included the selection of hairstyles, hair colors, eye colors, and one of three predefined outfits (casual, sporty, or elegant). This structured yet flexible customization was intended to enhance identification with the avatar while maintaining visual standardization across users.

After registration, participants were required to complete two exercises, *Dialogue with the Inner Critic* and *Approach–Avoidance Task*. In the *Dialogue with the Inner Critic* exercise, users first selected a Critic avatar from a set of six virtual agents (three female and three male) differing in skin tone and characterized by standardized yet elegant clothing. Participants then chose the domain of self-criticism to be addressed, including one Moral domain (criticisms related to one's moral behavior) and three Non-Moral domains (interpersonal relationships, personal characteristics, and physical appearance) according to the theoretical framework developed by Zaccari & Mancini (2026).

Based on the theoretical framework of Miceli and Castelfranchi (2018), different critical statements were collected through focus groups involving therapists. By recording the internal criticisms reported by patients during clinical practice, the researchers categorized the domains of criticism based on the presence of moral elements (deontological or altruistic) and/or non-moral content (such as abilities,

https://doi.org/10.53240/2026TOPIC5N2.111

physical appearance, etc.) (Zaccari & Mancini, 2026).

The Moral domain included a total of five critical statements. Two criticisms referred to perceived inadequacy with respect to deontological moral standards and concerns: (1) perversion and disgust toward immoral thoughts and (2) selfish behaviors. The remaining three moral criticisms referred to perceived inadequacy with respect to moral standards grounded in altruistic criteria and included: (3) the altruistic criterion related to fortune and misfortune; (4) lack of help and emotional closeness toward others; and (5) causing harm to others for personal purposes.

The Non-Moral domains comprised a total of six critical statements referring to perceived inadequacy, incompetence, or incapacity with respect to non-moral standards. Two criticisms concerned the criterion of relations with others, namely Inclusion and Agreeableness; two referred to personal characteristics, namely Abilities, Competence and Intelligence and Personality and Character; and two concerned physical appearance, namely the aesthetic criterion and physical attractiveness.

Once the domain and the specific criticism to be addressed had been selected, the user listened to the chosen Critic avatar replicating the selected criticism. Synchronized facial expressions and movements accompanied the criticism increasing the ecological validity of the interaction between participants and the Critic avatar. While the verbal content of the critical statements was developed by the research team to ensure coherence with the underlying theoretical model and clinical relevance, vocal features were generated using artificial intelligence. Participants could replay the criticism multiple times before responding.

After listening to the criticism, participants completed the manipulation check assessing state levels of emotions, self-criticism, self-compassion and self-esteem.

Furthermore, for each criticism within each domain, three predefined responses were provided. Every response was constructed by the research team according to the principles of Socratic questioning: logical-empirical (Is that true?/ What evidence do you have?), pragmatic (Does it work?), economic (Is it worth it?), moral (Is it fair?), in line with the contents of the chosen criticism. E.g. "How could you do that! It's not right. You went too far.", this is an example of a moral criticism, and there follows one of the three Socratic responses that could have been chosen by participants: "The fact that I made a mistake doesn't mean that I am a bad person (logical-empirical: Is that true?). Speaking to me this way changes nothing (pragmatic: Does it work?): it's better to learn from the experience and focus on my values instead of punishing myself (economic: Is it worth it?). I deserve understanding, too (moral: Is it fair?)."

Participants then selected one of the three responses or freely generated a personal response to the criticism, recording it aloud within the app. During the final phase, the participant's avatar and the Critic avatar were simultaneously displayed, and users listened to their recorded response as if it was directly addressing the Critic. All recordings were stored and could be reused across sessions. At the end of the Dialogue with the Inner Critic, the manipulation check was re-assessed.

The second exercise was based on the Approach–Avoidance Task (AAT), a behavioral paradigm commonly used to assess implicit approach and avoidance tendencies toward emotional stimuli. Although presented as a game-like interaction between the avatar and criticism-related adjectives from moral and non-moral domains, the task served both as a measure of implicit behavioral tendencies and as

https://doi.org/10.53240/2026TOPIC5N2.111

a reinforcement of the preceding Dialogue with the Inner Critic exercise.

Instructions for participants were as follows: first, participants selected their preferred focus area (moral vs. non-moral domains). They were then presented with a list of 40 domain-specific adjectives (20 positive and 20 negative). The task instructions were: "On the left side of the screen, you will see your avatar" (created during the registration phase). "Your task is to move the avatar along the bar, freely choosing whether to move closer to or further away from the words presented. Once you release the slider, you will not be able to change the avatar's position. Make sure to click only once. When you are ready, click the button below to start."

The task was administered in landscape mode on a mobile device. An avatar positioned at the center of a horizontal bar (visual analogue scale; 0 = left endpoint, 1 = right endpoint) was controlled by dragging a cursor. Adjectives were presented sequentially on the right side of the screen.

Participants moved the avatar along the horizontal bar to either approach the stimulus (rightward movement) or avoid it (leftward movement). Upon release of the cursor, the avatar automatically returned to the center and the next stimulus appeared. Reaction times (RTs) were defined as the interval between cursor selection and release. The task ended automatically after all stimuli had been presented.

Procedures

At baseline (T0) recruited participants completed the first survey, administered via Psytoolkit platform (Stoet, 2010; 2016), which included demographic information and the set of questionnaires (OCI-R and SIAS). Then they downloaded the SOCRATE App on their personal smartphone and used it daily for 15 days. Participants were required to use the app at least once daily. To ensure compliance and app usage, daily standardized email reminders were sent to all participants. Although participants were instructed to use the app once a day, they were free to perform both exercises, or just one of them, multiple times a day as needed. After 15 days of app usage, they completed the post-test survey (T1) including the same set of questionnaires. The overall study procedure is graphically represented in Figure 1. This repeated-measures design was selected to examine changes in symptom domains associated with self-critical processes, as well as to assess the feasibility of repeated digital assessments within a pilot framework.

Data Analysis

Based on the description provided in the previous section, the following measures and tasks were administered: Obsessive–Compulsive Inventory–Revised (OCI-R; Foa et al., 2002; Italian version by Marchetti et al., 2010), Social Interaction Anxiety Scale (SIAS; Mattick & Clarke, 1998; Italian validation by Sica et al., 2007), Manipulation Check, and the two tasks, Dialogue with the Inner Critic and Approach–Avoidance. Although these data were collected, they are not the focus of the hypotheses for the present pilot study; instead, they will be discussed in future papers regarding the testing and evaluation of the SOCRATE app. The primary hypothesis of this pilot study is to assess the feasibility of SOCRATE by comparing participants' scores on measures related to OCD and Social Anxiety symptoms before and after using the app.

Statistical analyses were conducted using Jamovi software (version 2.6.44). Based on a previously established cutoff (x>21) adopted in prior research on OCI-R (see Jensen, 2003), Participants were divided into two subgroups

https://doi.org/10.53240/2026TOPIC5N2.111

Figure 1: Schematic flowchart of the experimental procedure and study timeline.

based on clinical severity: a low-trait group (n = 11) and an high-trait group (n = 5). To examine the effects of app usage, a Mixed-design Analysis of Variance (ANOVA) was performed with Group (low vs high OCD scores) as a between-subjects factor and Time (T0 vs. T1) as a within-subjects factor on OCI-R scores. A one-way repeated measures ANOVA was conducted on SIAS scores to evaluate changed across Time (T0 vs. T1).

Results

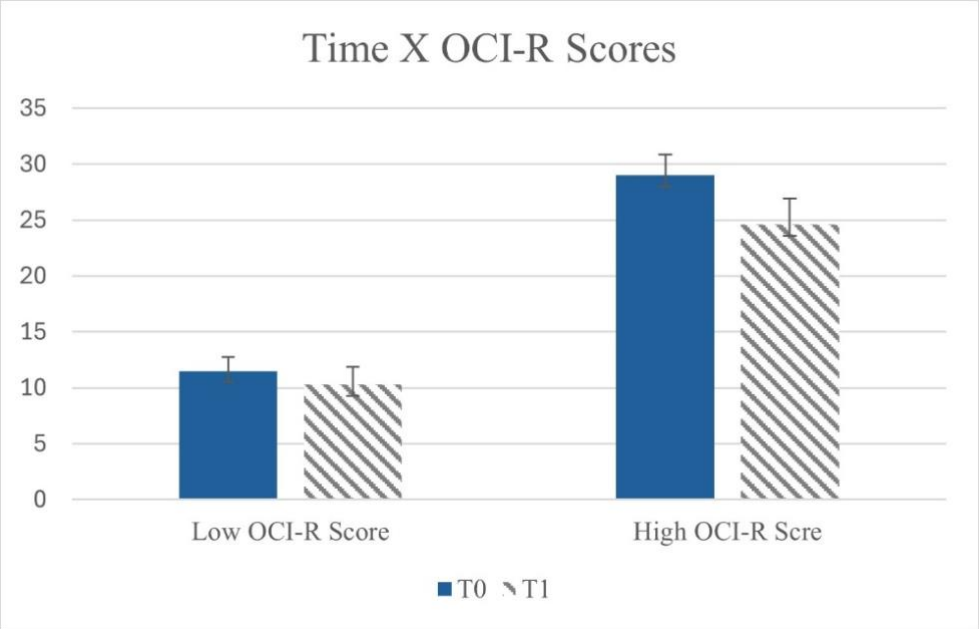
Significant effects emerged on OCI-R scores. A significant main effect of Time was found, indicating that app use was associated with an overall reduction in OCD symptoms regardless of groups, $F(1,14) = 19.03$, $p < .001$ (pre: M = 20.3; post: M = 17.4). A significant main effect of Group also emerged, $F(1,14) = 42.5$, $p < .001$, with higher OCI-R scores in the high OCD-traits participants (M = 26.8) compared to the normal group (M = 10.9). A significant Time × Group interaction was observed (see Figure 2). Post-hoc comparisons showed a significant reduction in OCD symptoms over time within the high OCI-R scores group (T0: M = 29.0; T1: M = 24.6; $p = .005$), although their scores remained significantly higher than those of the low OCI-R scores group (T0: low M = 11.5 vs. high OCI-R scores M = 29.0, $p < .001$; T1: low M = 10.3 vs. high OCI-R scores M = 24.6, $p < .001$). No significant pre–post changes were observed within the normal group. Regarding Social Anxiety, the repeated measures ANOVA showed no significant main effect of Time on SIAS scores ($F(1,15) = 0.01$; $p = .905$; T0: M = 30.8; T1: M = 30.6).

Discussion

The present findings provide preliminary support for the feasibility of engaging self-critical processes through a digitally structured Socratic Dialogue (SD) intervention. Repeated engagement with SOCRATE was associated with a reduction in Obsessive–Compulsive symptomatology in individuals presenting high OC scores. These results align with the evidence that individuals with OCD traits often exhibit chronic self-blame, excessive concern over mistakes, and rigid internal standards, which may result in high self-critical evaluations of the self (Golestaneh et al., 2017). This is consistent with the widely cited model on Self-Criticism (SC) by Gilbert et al. (2004), that distinguishes between two forms of SC: (a) SC characterized by feelings of inadequacy and inferiority, and (b) SC marked by self-directed hatred, disgust, and hostility. Although both forms lead to a range of adverse life outcomes (Zuroff et al., 1994; Gilbert et al., 2006), evidence suggests that in Eating Disorders (EDs) and Social Anxiety, SC tends to operate within non-moral domains, centering on perceived incompetence, unworthiness, inferiority, or defects in physical appearance and

https://doi.org/10.53240/2026TOPIC5N2.111

Figure 2: Pre- to post-intervention changes (T0 vs. T1) in the Obsessive-Compulsive Inventory-Revised (OCI-R) total scores, based on subclinical severity groups.



performance (Fairburn et al., 2003; Werner et al., 2011). Conversely, in Obsessive-Compulsive Disorder (OCD) and Depression, SC is often morally oriented and intertwined with excessive responsibility, self-blame, and deontological guilt (Mancini & Gangemi, 2021; Mancini et al., 2021). Contrary to our initial expectations, no significant reductions in social anxiety symptoms were observed following the use of the application. The most plausible explanation for this finding is likely the limited sample size. However, as this research represents a pilot study, we remain confident that future investigations will more effectively explore the relationship between SD and social anxiety through digital tools such as SOCRATE.

On this basis, SOCRATE could represent an innovative complementary tool to psychotherapy; furthermore, the results are consistent with the hypothesis that self-critical processes may be responsive to structured dialogical engagement. Importantly, the rule-based architecture of the application ensured procedural coherence and containment of the dialogical exchange supporting the feasibility of translating Socratic questioning into a controlled digital format. However, some limitations must be acknowledged. Specifically, the small sample size and the unequal gender distribution (predominantly female) limit the generalizability of the findings. Future research should consider employing larger and more demographically balanced samples to ensure greater robustness and replicability of the results. Future studies should also include follow-up assessments, essential to define the durability and consolidation of observed changes. Mechanism-focused studies are also warranted to directly assess whether reductions in obsessive–compulsive symptoms are mediated by modifications in moral self-evaluation, inferential rigidity, or the intensity and structure of the inner critical dialogue. From a clinical perspective, based on the emerging results,

https://doi.org/10.53240/2026TOPIC5N2.111

testing SOCRATE within diagnosed OCD populations and transdiagnostic samples characterized by elevated SC will clarify its generalizability and boundary conditions. Finally, future studies could explore its integration within care models evaluating whether structured digital dialogue enhances engagement, accelerates symptom reduction, or facilitates therapeutic alliance. These preliminary findings suggest that targeting the dialogical structure of the inner critic may represent a theoretically grounded and clinically scalable strategy for intervention.

Conclusions

In conclusion, this pilot study provides preliminary evidence that self-critical processes can be engaged through a structured, rule-based digital Socratic Dialogue (SD) intervention. Although in this preliminary study significant results have been observed only for obsessive–compulsive symptomatology, the findings support the broader feasibility of targeting the dialogical architecture of the inner critic within a transdiagnostic framework. By operationalizing SD within a rule-governed inferential sequence, SOCRATE illustrates how core CBT mechanisms can be translated into a scalable and clinically coherent digital format. These results contribute to the ongoing shift in digital mental health research from symptom-focused content delivery toward mechanism-driven intervention design. Targeting the dialogical structure of SC may represent a theoretically grounded and scalable strategy for enhancing psychological intervention in both clinical and subclinical populations.

Ringraziamenti

This study was funded by the Italian Ministry of University and Research (MUR) within the PRIN 2022 project SOCRATE- Self-criticism On Chair: Realizing an Application for Therapeutic Exercises (Principal Investigator: Prof. Francesca D'Olimpio, proj. #20222B4ZHF, CUP: B53D23014110006; CUP MASTER B53D2301411 0006).

Bibliografia

Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects.* New York: Harper & Row.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders.* International Universities Press.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression.* New York: Guilford Press.

Blease, C., & Torous, J. (2023). ChatGPT and mental healthcare: balancing benefits with risks of harms. *BMJ Mental Health, 26*(1), e300884. https://doi.org/10.1136/bmjment-2023-300884

Cerea, S., Ghisi, M., Bottesi, G., Carraro, E., Broggio, D., & Doron, G. (2020). Reaching reliable change using short, daily, cognitive training exercises delivered on a mobile application: The case of Relationship Obsessive Compulsive Disorder (ROCD) symptoms and cognitions in a subclinical cohort. *Journal of Affective Disorders, 276,* 775–787. https://doi.org/10.1016/j.jad.2020.07.043

Clark, G. I., & Egan, S. J. (2015). The Socratic Method in Cognitive Behavioural Therapy: A Narrative review. *Cognitive Therapy and Research, 39*(6), 863–879. https://doi.org/10.1007/s10608-015-9707-3

Fairburn, C. G., Cooper, Z., & Shafran, R. (2003). Cognitive behaviour therapy for eating disorders: a "transdiagnostic" theory and treatment. *Behaviour Research and Therapy, 41*(5), 509–528. https://doi.org/10.1016/s0005-7967(02)00088-8

Foa, E. B., Huppert, J. D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., & Salkovskis, P. M. (2002). The Obsessive-Compulsive Inventory: development and validation of a short version. *Psychological assessment, 14*(4), 485–496.

Gilbert, P. (2000). The relationship of shame, social anxiety and depression: The role of the evaluation of social rank. *Clinical Psychology & Psychotherapy, 7*(3), 174–189. https://doi.org/10.1002/1099-0879(200007)7:3<174::AID-CPP236>3.0.CO;2-U

Gilbert, P., Clarke, M., Hempel, S., Miles, J., & Irons, C. (2004). Criticizing and reassuring oneself: An exploration of forms, styles and reasons in female students. *British Journal of Clinical Psychology, 43*(1), 31–50. https://doi.org/10.1348/014466504772812959

Gilbert, P. (2005). *Compassion: Conceptualizations, research and use in psychotherapy.* New york, NY: Routledge.

Gilbert, P., & Irons, C. (2005). *Focused therapies and compassionate mind training for shame and self-attacking.* In P. Gilbert (Ed.), Compassion: Conceptualisations, research and use in psychotherapy (pp. 263–325). Routledge.

Gilbert, P., Durrant, R., & McEwan, K. (2006). Investigating relationships between perfectionism, forms and functions of self-criticism, and sensitivity to put-down. *Personality and Individual Differences, 41*(7), 1299–1308. https://doi.org/10.1016/j.paid.2006.05.004

Gilbert, P. (2014). The origins and nature of compassion focused therapy. *British journal of clinical psychology, 53*(1), 6-41.

Golestaneh, S. M., Dehghani, F., & Hoseini, F. S. (2017). Comparison of Self-Criticism in Obsessive-Compulsive, Major depression patients and normal people. *Jentashapir Journal of Health Research, 8*(3). https://doi.org/10.5812/jjhr.62161

Iftikhar, Z., Xiao, A., Ransom, S., Huang, J., & Suresh, H. (2025). How LLM Counselors Violate ethical Standards in Mental Health Practice: A Practitioner-Informed Framework. *Proceedings of the AAAI/ACM Conference on AI Ethics and Society, 8*(2), 1311–1323. https://doi.org/10.1609/aies.v8i2.36632

Im, C., & Woo, M. (2025). Clinical efficacy, therapeutic mechanisms, and implementation features of Cognitive Behavioral Therapy–Based Chatbots for Depression and Anxiety: Narrative review. *JMIR Mental Health, 12*, e78340. https://doi.org/10.2196/78340

Jensen, M. (2003). Interpretation of visual analog scale ratings and change scores: a reanalysis of two clinical trials of postoperative pain. *Journal of Pain, 4*(7), 407–414. https://doi.org/10.1016/s1526-5900(03)00716-8

Kazantzis, N., Fairburn, C. G., Padesky, C. A., Reinecke, M., & Teesson, M. (2014). Unresolved issues regarding the research and practice of cognitive behavior therapy: The case of guided discovery using Socratic questioning. *Behaviour Change, 31*(1), 1–17. https://doi.org/10.1017/bec.2013.29

Leary, M. R., Tambor, E. S., Terdal, S. K., & Downs, D. L. (1995). Self-esteem as an interpersonal monitor: The sociometer hypothesis. *Journal of Personality and Social Psychology, 68*(3), 518–530. https://doi.org/10.1037/0022-3514.68.3.518

Longe, O., Maratos, F. A., Gilbert, P., Evans, G., Volker, F., Rockliff, H., & Rippon, G. (2010). Having a word with yourself: Neural correlates of self-criticism and self-reassurance. *NeuroImage, 49*(2), 1849–1856. https://doi.org/10.1016/j.neuroimage.2009.09.019

Löw, C. A., Schauenburg, H., & Dinger, U. (2020). Self-criticism and psychotherapy outcome: A systematic review and meta-analysis. *Clinical Psychology Review, 75*, 101808. https://doi.org/10.1016/j.cpr.2019.101808

Mancini F. *The obsessive mind: understanding and treating obsessive-compulsive disorder*. New York: Routledge (2019).

Mancini F, Rogier G, Gangemi A. The distinction between altruistic and deontological guilt feelings: insights from psychopathology. In: Cardella V, Gangemi A, editors. *Psychopathology and philosophy of mind.* London: Routledge (2021). p. 21.

Mancini, F., & Gangemi, A. (2021). Deontological and Altruistic Guilt Feelings: a dualistic thesis. *Frontiers in Psychology, 12,* 651937. https://doi.org/10.3389/fpsyg.2021.651937

Marchetti, I., Chiri, L. R., Ghisi, M., & Sica, C. (2010). Obsessive-Compulsive Inventory-Revised (OCI-R): presentazione e indicazioni di utilizzo nel contesto italiano. *ArTS Archivio Della Ricerca Di Trieste (University of Trieste https://www.units.it/), 16*(1), 69–84. http://hdl.handle.net/11368/2935646

Mattick, R. P., & Clarke, J. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy, 36*(4), 455–470. https://doi.org/10.1016/s0005-7967(97)10031-6

Miceli, M., & Castelfranchi, C. (2018). Reconsidering the differences between shame and guilt. *Europe's Journal of Psychology, 14*(3), 710–733. https://doi.org/10.5964/ejop.v14i3.1564

Nagy, L. M., Shanahan, M. L., & Baer, R. A. (2021). An experimental investigation of the effects of self-criticism and self-compassion on implicit associations with non-suicidal self-injury. *Behaviour Research and Therapy, 139*, 103819. https://doi.org/10.1016/j.brat.2021.103819

Neff, K. D. (2003). Self-compassion: An alternative conceptualization of a healthy attitude toward oneself. *Self and Identity, 2*(2), 85–101. https://doi.org/10.1080/15298860309032

Neff, K. D., & Vonk, R. (2009). Self-compassion versus global self-esteem: Two different ways of relating to oneself. *Journal of Personality, 77*(1), 23–50. https://doi.org/10.1111/j.1467-6494.2008.00537.x

Overholser, J. C. (2010). Psychotherapy according to the Socratic method: integrating ancient philosophy with contemporary cognitive therapy. *Journal of Cognitive Psychotherapy, 24*(4), 354–363. https://doi.org/10.1891/0889-8391.24.4.354

Overholser, J. C. (2013). Guided discovery. *Journal of Contemporary Psychotherapy, 43*(2), 73–82. https://doi.org/10.1007/s10879-012-9229-1

Overholser, James. *The Socratic Method of Psychotherapy*, New York Chichester, West Sussex: Columbia University Press, 2018. https://doi.org/10.7312/over18328

https://doi.org/10.53240/2026TOPIC5N2.111

Padesky, C. A. (1993, September). *Socratic questioning: Changing minds or guiding discovery?* Keynote address presented at the European Congress of Behavioural and Cognitive Therapies, London, United Kingdom. http://www.padesky.com/clinicalcorner/

Pekin, Z., & Güme, S. (2025). Self-Criticism: Conceptualization, assessment and interventions. *Psikiyatride Guncel Yaklasimlar - Current Approaches in Psychiatry, 17*(1), 107–123. https://doi.org/10.18863/pgy.1455185

Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001b). Measuring Global Self-Esteem: Construct Validation of a Single-Item Measure and the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin, 27*(2), 151–161. https://doi.org/10.1177/0146167201272002

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton University Press.

Saliani, A. M., Perdighe, C., Zaccari, V., Luppino, O. I., Mancini, A., Tenore, K., & Mancini, F. (2024). Treating Guilt-Inducing Self-Talk in Ocd with Dramatized Socratic Dialogue: A Step by Step Intervention. *Clinical neuropsychiatry, 21*(1), 63–78. https://doi.org/10.36131/cnfioritieditore20230601 04

Sica, C., Musoni, I., Chiri, L. R., Bisi, B., Lolli, V., & Sighinolfi, C. (2007). Social Phobia Scale (SPS) e Social Interaction Anxiety Scale (SIAS): traduzione e adattamento italiano. *Bollettino di Psicologia Applicata, 252*, 59–71.

Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods, 42*(4), 1096–1104. https://doi.org/10.3758/brm.42.4.1096

Stoet, G. (2016). PsyToolkit. *Teaching of Psychology, 44*(1), 24–31. https://doi.org/10.1177/0098628316677643

The jamovi project (2025). jamovi (Version 2.6) [Computer Software]. Retrieved from https://www.jamovi.org

Werner, K. H., Jazaieri, H., Goldin, P. R., Ziv, M., Heimberg, R. G., & Gross, J. J. (2011). Self-compassion and social anxiety disorder. *Anxiety Stress & Coping, 25*(5), 543–558. https://doi.org/10.1080/10615806.2011.608842

Williams, B. M., & Levinson, C. A. (2022). A model of self-criticism as a transdiagnostic mechanism of eating disorder comorbidity: A review. *New Ideas in Psychology, 66*, 100949. https://doi.org/10.1016/j.newideapsych.2022.100949

Zaccari, V. & Mancini, F. (2026, January 23) *Convegno d'Inverno per il Cognitivismo Clinico 2026: Self-criticism: Dai modelli teorici alla pratica clinica.* Scuole di Psicoterapia Cognitiva. https://eventi.apc.it/Home/NewCoursePage/convegno-dinverno-per-il-cognitivismo-clinico-2026-self-criticism-dai-modelli-teorici-alla-pratica-clinica

Zaccari, V., Mancini, F., & Rogier, G. (2024). State of the art of the literature on definitions of self-criticism: a meta-review. *Frontiers in Psychiatry, 15*, 1239696. https://doi.org/10.3389/fpsyt.2024.1239696

Zuroff, D. C., Koestner, R., & Powers, T. A. (1994). Self-criticism at age 12: A longitudinal study of adjustment. *Cognitive Therapy and Research, 18*(4), 367–385. https://doi.org/10.1007/bf02357511