RESEARCH ARTICLE

Profiling adolescents' vulnerability to racial misinformation: An hybrid intervention aimed at promoting mediated intergroup contact

Francesca D'Errico[1] ⓘ | **Giuseppe Corbelli**[2] ⓘ |
Paolo Giovanni Cicirelli[1] ⓘ | **Carmela Sportelli**[1] ⓘ |
Marinella Paciello[2] ⓘ

[1]For.Psi.Com, Università degli Studi di Bari "Aldo Moro", Bari, Italy

[2]Faculty of Psychology, Uninettuno Telematic International University, Rome, Italy

Correspondence
Paolo Giovanni Cicirelli, For.Psi.Com,
Università degli Studi di Bari "Aldo Moro",
Bari, Italy.
Email: paolo.cicirelli@uniba.it

Funding information
Fondazione Compagnia di San Paolo,
Challenge for Europe' European Call,
Grant/Award Number: B99C20000640007

Abstract

This study addresses the issue of adolescents' susceptibility to racial misinformation, testing a socio-analytical intervention within an educational community through the inducement of analytical processing of misleading news on one side, and mediated contact on the other. Rolling Minds web app has been designed to implement a hybrid schools intervention thanks to which classmates engage with conversational agents, guiding them in deconstructing racial stereotypes, reframing misleading narratives and empathizing with immigrants' point of view. All the intervention activities aim to enhance intergroup contact. For this purpose, in this study, which involved 208 young participants (mean age 14.65; SD = 0.74), we pursued two research objectives. Firstly, by employing a person-centred approach, our first research question (RQ1) is to identify subgroups of adolescents based on their propensity to engage in analytical reasoning, stereotypical beliefs and self-transcendence values, finding different vulnerability clusters to racial

misinformation. In relation to the emerging profiles, their reactions to the misleading news were initially observed (ie, sharing and fact-checking). Secondly, regarding the second research question (RQ2), through path analysis we aim to understand whether and how adolescents characterized by different profiles vary in performing activities of socio-analytical intervention aimed at enhancing contact intentions. These results can guide the development of tailored hybrid educational strategies to engage young people in approaching racial online misinformation in a more reflective and unbiased manner, considering the adolescents' individual differences in vulnerability to racial hoaxes.

1 | INTRODUCTION

The spread of misinformation is currently a phenomenon of great interest to researchers, educators, teachers and psychologists, as it significantly impacts individuals and communities in terms of health, civic and political behaviours. This contribution aims to understand how hybrid technologies applied in an educational community can promote intergroup contact to prevent anti-immigrant prejudice, which is often linked to racial misinformation (Papapicco, Lamanna, & D'Errico, 2022; Wright, Brinklow-Vaughn, Johannes, & Rodriguez, 2021). In particular, to promote intergroup contact, we focus on the psycho-social intervention regarding a specific form of misinformation associated with prejudice and stereotypes, known as racial hoaxes, which adolescents may encounter on social networks (Papapicco et al., 2022). Racial hoaxes are defined as communicative acts based on distorted (but not necessarily false) information as a threat to health or safety, where the perpetrator is described in terms of ethnicity, nationality or religion (Cerase & Santoro, 2018; D'Errico, Cicirelli, Corbelli, & Paciello, 2023).

The stereotypical and schematic portrayal of immigrants can be an effective means of spreading biased information, reinforcing anti-immigrant attitudes and discrimination. This effect is evident in adults (Wright et al., 2021) as well as in young people (Wright & Duong, 2021).

To address this issue, the current study focuses on a socio-psychological intervention targeted at young students through a conversational web app, focusing on racial hoaxes.

Specifically, the study examines the effectiveness of a hybrid web app named 'Rolling Minds', used within educational communities, by taking into account the users' motivational and cognitive profiles. In this app, each student can respond to psychological measures by interacting with a virtuous and collaborative artificial agent (a peer named Sofia), thus allowing for the user's profiling. Then, the artificial agent helps them identify potential racial stereotypes, reframe the misleading news with alternative sources and consider the immigrant's perspective in the fabricated story, using a simulated conversation through a chat. This online interaction is complemented by an important experiential aspect of countering misinformation, that is, the guided debriefing (Greene & Murphy, 2023), where participants discuss and apply their experiences in a real context with classmates, teachers and psychologists.

In this context, hybrid technology (Gatti & Procentese, 2020; Gatti, Procentese, & Schouten, 2023) serves as a tool to promote the so-called mediated intergroup contact, an indirect type of contact with a member of the

outgroup performed by means of different mass media formats such as newspapers, TV or also books (Banas, Bessarabova, & Massey, 2020; Goldman & Mutz, 2010; Vezzali, Hewstone, Capozza, Giovannini, & Wölfer, 2017; Visintin, Voci, Pagotto, & Hewstone, 2017). This intervention includes interactions with an outgroup member who shares the story from his perspective, thus offering a common and controlled experience with classmates. The conversational web app also employs the media bias approach, thus focusing on misleading news' stereotypes or on the description of the ingroup members unique point of view (Paul & Elder, 2006), by fostering analytical processing of adolescents who are, surprisingly, vulnerable to such news, including racially biased information (Faragó et al., 2024; Herrero-Diz, Conde-Jiménez, & Reyes-de-Cózar, 2021; Papapicco et al., 2022).

Starting from the theoretical and practical premises of this socio-analytical intervention, this contribution aims in particular to explore the interplay among cognitive and motivational dimensions in preventing racial misinformation and ethnic prejudice by examining the role of adolescents' differences in understanding the effectiveness of the intervention using the hybrid conversational web app. Specifically, by adopting a person-centred approach (Magnusson & Stattin, 1996), our first research question (RQ1) is to identify adolescents' profiles based on the interplay among analytical thinking and two motivational dimensions: self-transcendence values and pre-existing stereotypes, along with their initial behavioural reactions to misleading news (ie, credibility, sharing intention and fact-checking). These potential profiles will emerge from the integration of literature on credulity towards misinformation and intergroup contact, which consider analytical thinking, values and stereotypes as crucial factors.

Additionally, our second research question (RQ2) is to test if and how the adolescents' profiles emerging from the interplay among cognitive and motivational dimensions (collected in the first phase of the conversational web app) affect the socio-analytical intervention aimed at enhancing intergroup contact intentions.

1.1 | Countering racial hoaxes: The role of values, stereotypes and analytical thinking

Previous studies have offered various, mostly complementary, explanations for people's susceptibility to fake news. Primarily, the content of fake news often aligns with deeply held beliefs, triggering identity protective cognition. Specifically, under the motivated reasoning theory (Kahan, 2013; Kahan et al., 2017), individuals utilize their reasoning abilities to protect and maintain their ideologies, and existing ideas, rather than to acquire accurate information, and they are driven to shield their beliefs from conflicting evidence (Kunda, 1990; Nir, 2011). Consequently, their attitudes towards certain ideas significantly shape their intentions and actions, leading them to disproportionately believe content that is congruent with their views, as a form of identity-protective cognition (Kahan, 2013). One of the ways people adhere to their motivated reasoning is through stereotypes and prejudices: stereotypes create expectations about what a member of a group would do or say, which can serve as guidelines for behaviour and decision-making (Rahn, 1993), while prejudice is defined as a negative feeling towards a group based on a faulty generalization (Pettigrew, 1982). Misleading news often targets minority groups to confirm stereotypes, validate prejudice and encourage discrimination (Wright et al., 2021; Wright & Duong, 2021). However, the relationship between prejudice, stereotypes and the identification of misleading news seems to be bilateral: not only can fake news elicit prejudices and stereotypes, but according to research on confirmation bias, people are also likely to interpret new information in line with their pre-existing knowledge (Nickerson, 1998). Wright and Duong (2021) underscore the relationship between the inability to identify fake news and attitudes towards the minority group targeted by such news. For instance, they find that the inability to discern COVID-19 fake news is linked with anti-Asian American sentiment and xenophobia. As motivational determinants, values can contribute to the credibility and sharing of fake news. People are motivated to protect their values and worldviews (Maio & Olson, 1998), which results in more positive attitudes towards targets who uphold these values (Rosenblatt, Greenberg, Solomon, Pyszczynski, & Lyon, 1989). D'Errico, Corbelli, Papapicco, and Paciello (2022) highlight how, depending on the moral content that misleads in news, during adolescence, conservation values and self-transcendence can exert opposing influences on the sharing of misleading news. Conservation refers to those values (conformity, tradition and security) fostering internal cohesion within a group, while self-transcendence refers to those

values (benevolence and universalism) aiming to promote prosocial bonds among individuals, even extending to those outside the group. Self-transcendence seems to discourage the sharing of misleading news related to violations of formalized norms (regulations), group conventions and respect for authority. Additionally, higher levels of conservation values promote the dissemination of misleading news about the violations of rules governing behaviour within a given society, irrespective of the cognitive assessment of the news' truthfulness.

Moreover, prior research suggests that when individuals are exposed to fake news aligned with their pre-existing beliefs or values, they tend to be less motivated to engage in analytical reflection regarding the news article itself (Allcott & Gentzkow, 2017; Taber & Lodge, 2006). Recent research highlights the role of a lack of analytical thinking in determining individual susceptibility to fake news. Pennycook and Rand (2019) found that those who exhibited high levels of analytical reasoning skills were better equipped to distinguish between fake and real news headlines, even after controlling for political ideology. Lutzke, Drummond, Slovic, and Árvai (2019) emphasize the importance of not regarding analytical thinking skills merely as pre-existing abilities but instead advocate for interventions aimed at priming analytical thinking in participants, thereby preventing the spread of misinformation. These aspects can underlie individuals' vulnerability profiles to racial misinformation and, above all, play a potential role in their interaction with prejudice and intergroup contact.

1.2 | Fostering intergroup contact intention: The impact of analytical thinking, values and stereotypes

The Contact Hypothesis (Allport, 1954) posits that frequent and positive interactions with outgroup members can enhance intergroup relations and reduce ethnic prejudice towards the outgroup. Studies focusing on prejudice prevention and the promotion of intergroup contact have also highlighted the significance of analytical thinking and motivated cognition, which takes into account both prejudice and values (Kahan, 2013), as two crucial dimensions. Numerous studies have revealed the relationship between analytical thinking and other cognitive skills with ethnic prejudice (Blanchar & Sparkman, 2020; Yilmaz, Karadöller, & Sofuoglu, 2016) and intergroup contact. Individuals with higher levels of analytical thinking tend to exhibit lower levels of prejudice and a greater inclination towards intergroup contact (Hurtado, 2005). In the context of intergroup contact, motivated reasoning, as the motivation to interpret social information in accordance with prior knowledge and consistent expectations, emerges again as a relevant consideration (Kahan, 2013; Kunda, 1990; Pereira, Harris, & Van Bavel, 2023). For instance, cognitive prejudice is negatively associated with positive intergroup contact across various contexts, including women's stereotypes about men (Stephan, Stephan, Demitrakis, Yamada, & Clason, 2000) and Whites' stereotypes towards Black people (Corenblum & Stephan, 2001; Stephan et al., 2002; Tropp & Pettigrew, 2005). In addition, values centred around self-transcendence and a general focus on openness to others and diversity play crucial roles in discouraging prejudice (Souchon, Maio, Hanel, & Bardin, 2017) and promoting intergroup contact with out-groups (Barni, Cavazza, Russo, Vieno, & Roccato, 2020; Sagiv & Schwartz, 1995). When analysing the literature on racial misinformation prebunking, and that relating to the promotion of intergroup contact, it seems important to take into consideration the cognitive factors and socio-motivated cognition (ie, stereotypes and values; Kahan, 2013) to understand the effects of intervention in terms of intergroup contact mediated by the analysis induced by the Rolling Minds web app.

1.3 | The person-centred approach to intergroup contact promotion

Given the literature described, the general aim of the present study is to observe how cognitive and socio-motivational aspects (motivated reasoning; Kahan, 2013) can vary in their effect on the socio-analytical intervention aimed at preventing racial misinformation by enhancing contact intentions. The adolescents can be in fact differently inclined to intergroup contact (Barni et al., 2020; Hurtado, 2005; Sagiv & Schwartz, 1995) and can benefit differently from a hybrid mediated intervention based on a socio-cognitive approach (D'Errico et al., 2023).

This issue will be pursued through two research questions:

RQ1. (a) Can different cognitive-motivational profiles of adolescents be distinguished based on how analytical reasoning, stereotypes and self-transcendence interact? (b) What are the behavioural and evaluative reactions of these profiles after reading the misleading news (ie, credibility of misleading racial news, sharing and fact-checking intention)?

RQ2. If so, do these cognitive-motivational profiles respond differently to the proposed intervention based on interactive conversational technology?

Related to the first research question, a person-centred approach has been employed (Magnusson & Stattin, 1996) to explore the potential interrelations among cognitive and socio-motivational aspects (ie, the propensity towards analytical reasoning, self-transcendence and stereotypes) to identify groups of adolescents differently susceptible to the effects of racial disinformation. The person-oriented approach (Howard & Hoffman, 2018) highlights differences among individuals in how variables relate to each other, focusing on the identification of groups that function similarly within themselves but differently from each other. For instance, some adolescents may be inclined to adopt analytical reasoning, but have low levels of self-transcendence, while others may exhibit high levels of both analytical reasoning and self-transcendence.

Precise expectations regarding the number of adolescent groups were not established; however, specific cluster profiles were hypothesized. It was anticipated that two extreme groups of adolescents could be identified: one group characterized by low levels of analytical reasoning, high levels of stereotyping and low levels of self-transcendence values, representing a group less inclined towards intergroup contact and potentially more susceptible to racial misinformation; and another group with the opposite profile, exhibiting high levels of analytical reflection, a low level of stereotyping and high levels of self-transcendence values, representing a more resilient group towards racial misinformation and more prone to intergroup contact. Intermediate profiles were also anticipated, where cognitive and motivational dimensions may vary due to individual differences in cognitive maturation and value development. For example, one might expect a group of adolescents who are generally open to diversity and do not hold high levels of stereotypes, yet have not developed sufficient analytical reasoning skills to recognize the manipulations of racial misinformation.

This identification approach holds the potential to inform tailored interventions to mitigate the impact of misleading racial news, recognizing that a one-size-fits-all intervention may not be universally effective. Person-centred analytic models assume population heterogeneity in how predictors influence outcomes (Szebeni, Jasinskaja-Lahti, Lönnqvist, & Szabó, 2023). In this specific case, various groups may describe different levels of adolescents' readiness in interpreting racial hoaxes cautiously, their susceptibility to biased interpretations and, in turn, their propensity towards intergroup contact.

For these reasons, by integrating a person-oriented approach with a variable-oriented approach, the second research question aimed to examine the effects of the socio-analytical intervention through the web-app by considering the role of different cognitive motivational profiles among adolescents through the activities suggested by the conversational agents within the web app. This investigation is particularly useful in determining whether the analysis of racial hoaxes is effective for adolescents with varying vulnerability mindsets concerning racial misinformation. It is hypothesized that adolescents with high levels of analytical reasoning and self-transcendence values and low levels of stereotypes are better equipped to cope with racial misinformation and perform better during conversational web app interventions compared to adolescents with opposite profiles. This group is also expected to exhibit a higher intention for intergroup contact both before and after the intervention. Moreover, investigating the impact of the intervention on adolescents with less equipped profiles can provide more useful insights into how the intervention can be effective for those who need it most, as they are more vulnerable to misinformation and less open to intergroup contact. Overall, integrating the person-centred approach to identify different profiles of adolescents with the variable-centred approach to observe their responses to the intervention can contribute to the literature on designing personalized educational interventions, including those utilizing innovative and hybrid technology.

2 | METHODS

2.1 | The socio-analytical intervention through 'Rolling Minds'

Considering previous studies on interventions aimed at countering misinformation (Traberg, Roozenbeek, & van der Linden, 2022), it becomes evident how the reliance on analytical reflection can foster resilience against misleading news (Pennycook & Rand, 2019), as well as against stereotypes and other distorted beliefs associated with racial misinformation (D'Errico et al., 2023).

Moreover, the Paul and Elder intervention approach (2006), based on young people's media biases, underscores the necessity of further promoting analytical processing by encouraging the identification of the point of view from which the story is told, identifying which perspectives are denied or ignored and distinguishing between the factual basis of the story and the interpretation given to those facts.

Considering the discriminatory aspects of racial hoaxes, alongside technologies to combat misinformation, it proves useful to also take into account the potential influence of mediated contact, that is, contact with an outgroup member through various forms of media (Banas et al., 2020; Goldman & Mutz, 2010; Visintin et al., 2017), as a means to mitigate potential associated stereotypes and prejudice.

In particular, reading or listening to personal stories (Figenschou & Thorbjørnsrud, 2015; Pettigrew & Tropp, 2006) can reduce participants' biased beliefs by promoting contact intentions (Birtel et al., 2019; Paravati, Fitzgerald, Green, McAllister, & Moore, 2022).

In light of these considerations, this intervention aimed to induce analytical processing by enhancing the recognition of racial hoaxes biases (D'Errico et al., 2023; Paul & Elder, 2006), and humanizing the outgroup member through mediated contact by having participants read the immigrant point of view on the distorted news (Goldman & Mutz, 2010). Additionally, the mediated nature of this intervention prompted consideration of how digital tools and applications, especially those integrated into offline educational contexts and thus hybrid in nature (Gatti & Procentese, 2020), could serve as a means to combine the necessity of fostering socio-analytical thinking regarding misleading news while simultaneously facilitating mediated contact with the potential victim of the racial hoax. Thus, such conversational intervention allows for the presentation of the outgroup member's perspective (Said, of African ethnicity), by reflecting, engaging and chatting with a peer educator avatar belonging to the participant's ingroup (Sofia, of Caucasian ethnicity).

2.2 | Participants

The research focused on adolescents and collected data from two schools located in Southern Italy using a convenience sampling method. The total number of participants was 208, comprising 49 males, 158 females and 1 non-binary. All participants are of Italian origin, except for two adolescents of African origin. The ages of the sampled individuals ranged from 13 to 17 years, with an average age of 14.65 (SD = 0.74).

2.3 | Procedure

Two schools, including a classical high school and a vocational high school, were involved in the study, and no student declined to participate. Informed consent was obtained from the legal representatives of the participants since they were minors. In the informed consent, it was described to participants as 'an investigation on personal preferences in the consumption of online news'. We explained that the entire procedure would take place through a conversational web-app, involving interactions with avatars, and emphasized that the research would be non-evaluative and strictly anonymous.

The entire procedure consisted of two 1-h sessions, followed by a detailed debriefing session with psychologists and teachers where participants were thoroughly informed about the research goals and involved in a shared discussion related to racial misinformation issues. All procedures strictly adhered to the ethical principles outlined in the Helsinki Declaration and were in alignment with the ethical code of the Italian Association for Psychology (AIP). Additionally, both the research design and procedures received approval from the ethics committee at the university, with which two of the authors are affiliated (reference code: ET-22-01).

2.3.1 | Rolling Minds web app

The web app, developed in collaboration with BSD Design, employs a conversational logic and is visually presented as a chat (see Figure 1). Participants engage with various avatars (Sofia and her aunt, Susanna), with predesigned turn-taking questions and potential answers during different subsequent phases (Table 1). The web app incorporates both predetermined answers in the form of multiple-choice questions and open-ended questions. All participants completed the web app independently. To enhance engagement throughout the process, users receive badges as feedback for their efforts.

Initially, the user is introduced into a conversation with Sofia, a peer educator, who, getting to know the participant, poses several questions and shares information about herself (pre-assessment phase). After this first phase, a Racial Hoax (RH) is presented through the conversational approach to assess the participant's behavioural intentions and credibility. Specifically, a customized RH was created, involving a violation of the moral domain of Care (Graham et al., 2011), which is known to strongly activate individuals with high levels of self-transcendence as well as high levels of prejudice (Athota, Kearney, & Cocodia, 2015; Hadarics & Kende, 2018). In this scenario, the protagonist seems to attack a disabled person. The constructed news piece, following the literature (Cerase & Santoro, 2018), mimics the appearance of an Instagram post and incorporates the typical features of racial hoaxes: a fictitious newspaper with a plausible name ('BreakNotizie24'), uncertain sources (two passers-by) and the inclusion of stereotypical terms ('clandestine', 'illegal').
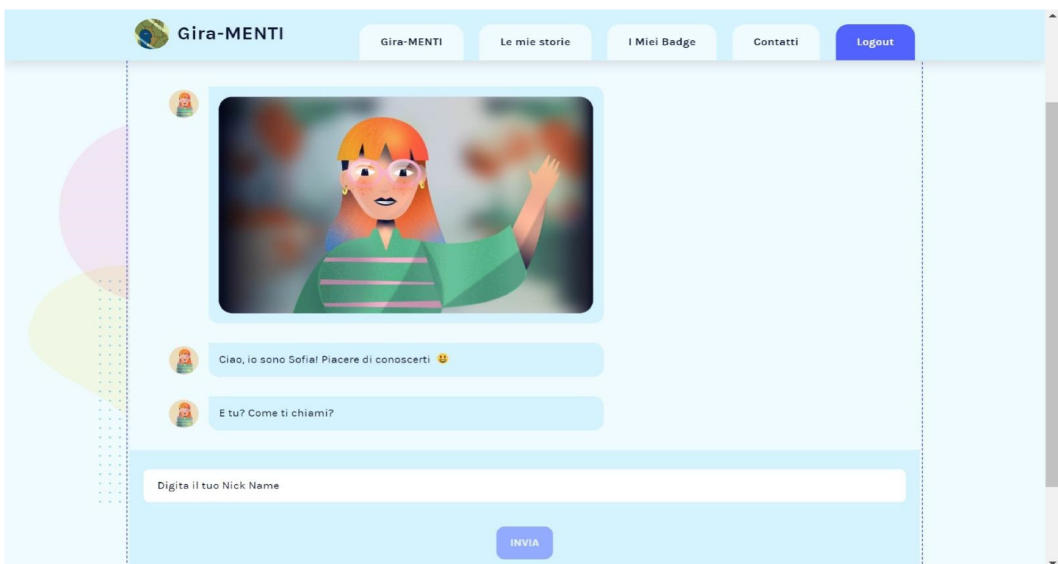

FIGURE 1 Rolling Minds interaction with Sofia.

TABLE 1 Web app phases description.

Phase	Avatar	Storytelling	Goals
Introduction + Nice to meet you, I'm Sofia	• Sofia	The user meets Sofia, a 14-year-old girl who wants to befriend the user after moving to a new city. While browsing social media, she runs into a wired news and asks the user what he/she thinks of it.	User profiling Pre-intervention measures (analytical thinking, self-transcendence values and stereotypes) are collected in a conversational way. Then, a racial hoax is administered, to collect the user's behavioural intentions and credibility.
Think about it auntie!	• Sofia • Susanna	Sofia's aunt, Susanna, has just shared the racial hoax presented in the first phase. Sofia, who is part of an activist group aimed at countering online misinformation and negative online behaviours, will ask the user for help. The user cooperates with Sofia to explain online misinformation, its characteristics and consequences to her aunt Susanna, who is not very familiar with social media.	Analytical reflection on media biases and mediated contact. Chatting with the peer educator Sofia, the user analyses and reflects on the racial hoax features. In particular, the user is involved in stereotype identification, source analysis, header recognition, separating fact from opinion and analysing points of view. Then, an alternative news is presented, highlighting the perspective of the outgroup member portrayed in the racial hoax (mediated contact). Finally, the user tries to rewrite the news in an analytical and unbiased way.
Conclusion— End of conversation	• Sofia • Susanna	After meeting again Sofia's 'boomer' aunt (aunt Susanna), Sofia and the user reflect about the media biases explained in the previous phase	Efficacy evaluation of intervention. Post-intervention measures are collected in a conversational way

During the conversational interaction, after the user and Sofia find out that the same RH has been shared by Sofia's aunt, Susanna, the two share a common goal of changing aunt Susanna's mind about sharing such information by explaining to her the distinct characteristics of RH ('analytical reading phase'). In this section, the implementation focused on Paul and Elder's procedure (2006), which had been previously tested in a prior study (D'Errico et al., 2023). Aided by the conversation with Sofia, the users are induced to focus on the media biases emerging from the misleading news. This includes recognizing the source, distinguishing between facts and judgements, recognizing existing stereotypes and fostering the ability to explore alternative interpretations of the events. After providing a simple explanation and some examples of each media bias, participants were asked to recognize and report each bias in the news they analysed.

After analysing the news with the assistance of the ingroup avatar Sofia, the mediated contact phase with the outgroup member Said started, adhering to the well-known optimal conditions (equal status among members, common and cooperative goals and support) outlined by Allport (1954). Specifically, the mediated contact occurs when Sofia, the peer educator, presents two alternative news stories (Figure 2), referred to as counterstories. In these counterstories, the immigrant character from the RH (initially described stereotypically and accused of aggression) explains his motivations: this narrative shift inverts the perspective, revealing that his true intent was to assist the disabled person, not to attack him. The primary goal of these counterstories was to humanize Said, as suggested by Paravati et al. (2022). This approach allowed the outgroup member to freely share his story and encouraged the user to consider his viewpoint on the event, thereby inducing the mediated contact phase. In the two counterstories, Said recounts his experiences as a scout and volunteer.

Following this, Sofia requests the user's assistance in rewriting the news, taking into account the two versions of the incident ('analytical rewriting phase'). The interaction with Sofia concludes with the completion of post-test measurements, including the assessment of the propensity for intergroup contact ('post-intervention phase'), and the invitation for the user to become an official debunker for the Rolling Mind group. This marks the end of the story within the web app. During the subsequent debriefing session with psychologists and teachers, the students have the opportunity to collectively discuss their mediated experiences with Sofia, her aunt Susanna (as avatars) and Said, the protagonist of the misleading racial news, debating the topics that emerged.

2.4 | Measures

Through different interactive modules and via the conversational storytelling approach, the Rolling Minds web app enables pre-assessment measurement, monitors performance throughout the intervention phase and allows for a subsequent assessment post-intervention.
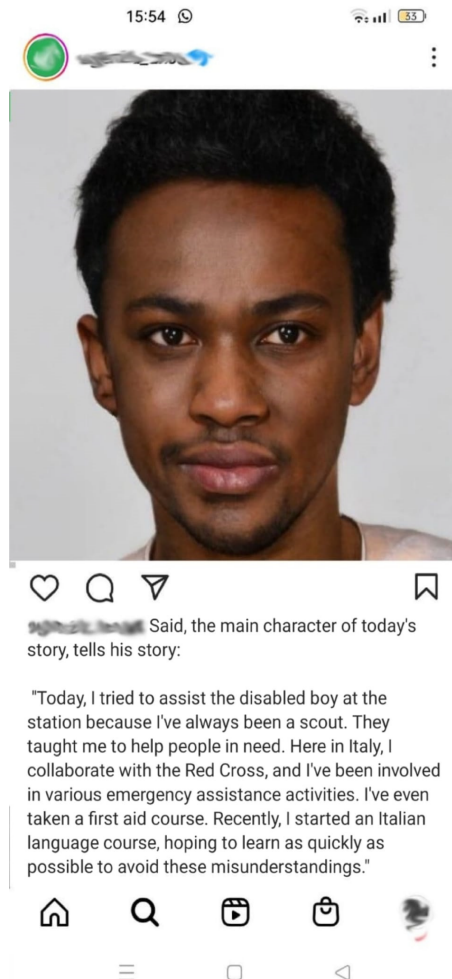

FIGURE 2 Example of the mediated contact in Rolling Minds: the news told by outgroup member (Said).

2.4.1 | Pre-assessment variables

Propensity to engage in analytical reasoning (Cognitive Reflection Test-2)

The two items previously found to be the most informative from the revised version of the Cognitive Reflection Test (CRT-2; Thomson & Oppenheimer, 2016) were specifically chosen to assess an adolescent's inclination to thoughtfully analyse a problem and resist the impulse to quickly choose the first plausible answer (Frederick, 2005), that is, the propensity to engage in analytical reasoning (PAR). The two items are, respectively, 'If you're running a race and you pass the person in second place, what place are you in?' where the intuitive but incorrect answer is 'first', while the correct answer is 'second'; and 'A farmer had 15 sheep, and all but 8 died. How many are left?' where the intuitive but incorrect answer is '7', while the correct answer is '8'. After the coding of the two items with the help of the R package reflectR (Corbelli, 2024), each correct response was assigned 1 point. The two scores were then averaged, resulting in final test scores ranging from 0 to 1. A score of 0 indicates the absence of correct reflective responses, while a score of 1 indicates two correct responses; an intermediate score of 0.5 indicates one correct response out of two.

Stereotypes

To assess the stereotype's scores among the participants regarding the African ethnic group, an evaluation was conducted using a 5-point Likert scale, ranging from 'Not at all agree' to 'Totally agree', on five adjectives inspired by the Stereotype Content Model (Cuddy et al., 2009). The five adjectives used were: bossy, intelligent, honest, affectionate and slacker. The scores for the three positive adjectives were reversed to create an index that measures increased negative stereotyping of the African group (Cronbach's alpha is 0.69).

Self-transcendence

Three items, assessed on a 6-point Likert scale, were adapted from the Portrait Values Questionnaire (Capanna, Vecchione, & Schwartz, 2005) to measure self-transcendence (eg, 'It's very important to him to help the people around him. He wants to care for their well-being'). Cronbach's alpha is 0.69.

Reactions to the racial hoax

After administering the news to the participants, their credibility and the behavioural intentions were measured using three different single-item measures. Adopting the procedure suggested by Pennycook and Rand (2021), credibility was assessed by asking participants to respond to the question 'How credible is the news you just read?' on a 5-point scale ranging from 1 ('Not at all') to 5 ('Totally'). For sharing intention, participants were asked 'Would you share it on social media?' and responded on a similar 5-point Likert scale ranging from 1 ('Definitely no') to 5 ('Definitely yes'). Fact-checking intention was evaluated by asking participants 'Would you look for other sources to compare if what the news says is true?' using the same 5-point Likert scale, adapting the procedure of Edgerly, Mourão, Thorson, and Tham (2020).

2.4.2 | Intervention variables

Analytical racial hoaxes reading and Racial hoaxes re-writing after mediated contact (ARHR and RHW)

To assess the individual ability for analytic reading and analytic rewriting of the news following the Mediated Contact phase, we utilized two previously tested indices (D'Errico et al., 2023). The first index, ARHR, measures the individual's capacity to recognize media biases. It consists of a series of tasks that participants completed, including multiple-choice questions and open-ended responses, which were subsequently coded by independent judges. The tasks included recognition of the source (Cohen's kappa = .98), recognition of facts (Cohen's kappa = .76),

identification of judgemental statements (Cohen's kappa = .84), identification of stereotypes in the Racial Hoax (RH) (Cohen's kappa = .91) and the ability to find alternative interpretations (Cohen's kappa = .85).

The second index, RHW, assesses the quality of the news rewritten by the participants after the mediated contact. It is based on all measures coded by independent judges, including the focus on the title rewritten (Cohen's kappa = .86), focus on the text rewritten (Cohen's kappa = .73), use of stereotypes in the title rewritten (Cohen's kappa = .81) and use of stereotypes in the text rewritten (Cohen's kappa = .84).

2.4.3 | Outcome variables

Intergroup contact intention

To assess the intention to engage with African people, two items were employed, rated on a 5-point Likert scale, with response options ranging from 'Not at all' to 'Completely'. These items were adapted from the work of Cameron, Rutland, and Brown (2007). Specifically, participants were asked: 'Could you spend your free time with an African boy?', 'Could you befriend him?' This short survey was administered both before and after the intervention.

2.5 | Planned analysis

In the initial phase, descriptive statistics and zero-order correlations for the variables of interest were analysed. Then, a hierarchical regression analysis was used to determine the incremental contribution of intervention variables in explaining post-intervention contact intentions, after accounting for baseline contact intentions, gender and age. With the aim of identifying distinct adolescent groups based on PAR, on stereotypical beliefs and self-transcendence, a hierarchical clustering procedure was applied using Ward's method with Euclidean distance. Hypothetical solutions for three, four and five clusters were explored. The number of clusters was preliminarily chosen by examining the dendrogram to observe the presence of a clear data structure, identifying significant partitions. The validity of the chosen clustering solution was assessed by choosing the solution maximizing Dunn's index and the mean Silhouette value for each observation, while minimizing connectivity index. The stability of the selected cluster solution was further evaluated by minimizing the Average Proportion of Non-overlap (APN) and the Average Distance between Means (ADM). The final cluster membership was determined through a non-hierarchical k-means clustering procedure, initialized from the centroids derived from the chosen hierarchical partition. Multivariate and univariate analyses of variance were thus conducted to assess the significance of the mean differences between groups for each of the three dimensions of interest.

To assess the specific modes of functioning of the identified adolescent groups regarding the variables concerning socio-analytical intervention mediated by conversational storytelling, the different profiles were examined. The analysis included credibility of the news, sharing intention, fact-checking intention, analytic performance in reading the fake news, performance in rewriting the fake news and contact intention. The significance of differences between the means was assessed through a multivariate analysis of variance, followed by univariate analysis.

On each group identified by cluster analysis, a paired sample test was preliminary performed with the aim of assessing the effectiveness of the social-analytical intervention in increasing adolescents' contact intentions towards an outgroup member. A path analysis model was fitted, focusing on the socio-cognitive processes manipulated by the web app that influence changes in contact intentions. This model incorporates controls for participant groupings derived from cluster analysis results, as well as for gender, age and the initial level of contact intentions. The indirect effects of belonging to different groups on post-intervention contact intentions were assessed through 95% confidence intervals via the bias-corrected bootstrap method with 5,000 samples (MacKinnon, Lockwood, & Williams, 2004).

All analyses were performed in the R environment (R Core Team, 2021) for statistical analysis using the packages lavaan (Rosseel, 2012), mvnormalTest (Zhang, Zhou, Shao, & Zhang, 2020), haven (Wickham, Miller, & Smith, 2023), cluster (Maechler, Rousseeuw, Struyf, Hubert, & Hornik, 2023), clValid (Brock, Pihur, Datta, & Datta, 2008), factoextra (Kassambara & Mundt, 2020), fpc (Hennig, 2023), psych (Revelle, 2024) and Hmisc (Harrell, 2024).

3 | RESULTS

3.1 | Preliminary analyses

Descriptive statistics and zero-order correlations for the dimensions of interests are shown in Tables 2 and 3, respectively. In particular, from Table 3 it can be seen that PAR inversely correlates with gender, suggesting that females might be less inclined towards analytical reasoning. It also shows positive links with both analytical racial hoax reading and racial hoax rewriting. Additionally, this propensity is positively related to self-transcendence. Meanwhile, self-transcendence itself has an inverse relationship with stereotypical beliefs and a positive association with sharing intention. Regarding fact-checking intention, it aligns positively with PAR and self-transcendence and inversely with stereotypical beliefs. Performance in analytical racial hoax reading also correlates positively with self-transcendence and fact-checking intention. Furthermore, racial hoax rewriting is positively linked with PAR and self-transcendence, and it also shows a connection with fact-checking intention. Contact intentions at T1 and T2 are correlated with each other and both inversely correlate with stereotypes scores, while showing positive correlations with PAR, self-transcendence and fact-checking intention. However, only contact intention at the first-time step is positively linked with sharing intention, while the same construct measured at the second-time step is positively correlated with racial hoax rewriting.

Then, a hierarchical regression analysis was conducted to assess whether the two variables related to the intervention (ie, ARHR and RHW) significantly added to the explanation of post-intervention contact intentions, after controlling for baseline contact intentions, gender and age. Missing data for this preliminary regression analysis were imputed using median values. The comparison with the null model indicated that including baseline contact intentions, gender and age as initial predictors improved the model fit ($R^2 = .36$; $F[3, 204] = 39.70$, $P < .001$) almost exclusively due to the influence of baseline contact intentions. Subsequent inclusion of the average of the two performance variables resulted in a significant increase in R^2 by 4.2% ($\Delta R^2 = .04$; $F[4, 203] = 14.10$, $P < .001$), confirming that the process performances variables ARHR and RHW contribute to the explanation of the outcome. The final model accounted for a total of 38.8% of the variance (adjusted R^2), while standardized beta coefficients showed

TABLE 2 Descriptive statistics for the considered variables.

	Mean	SD	Sk	K
Propensity to engage in analytical reasoning	0.55	0.38	−0.17	−1.25
Stereotypical belief	1.05	0.60	0.28	0.14
Self-transcendence	3.56	0.90	−0.64	0.28
Credibility	2.17	0.82	0.09	0.47
Sharing intention	1.91	1.13	−0.26	−0.77
Fact-checking intention	2.93	1.06	−0.91	0.23
Analytical racial hoax reading	0.00	1.00	−0.51	0.04
Racial hoax re-writing	0.00	1.00	−0.18	−0.72
Contact intention (T1)	2.69	0.94	−0.53	−0.05
Contact intention (T2)	2.85	1.04	−0.82	0.09

TABLE 3 Zero-order correlation matrix.

	1	2	3	4	5	6	7	8	9	10	11
1. Gender	-										
2. Age	−0.11	-									
3. PAR	−0.22**	0.00	-								
4. Ster.	−0.06	−0.01	0.02	-							
5. Self.Tr.	0.17*	−0.06	0.23**	−0.26***	-						
6. Credib.	0.10	−0.22**	0.03	−0.20**	0.13	-					
7. Share	0.05	−0.28***	−0.08	−0.15*	0.17*	0.28***	-				
8. FC	−0.07	0.04	0.18*	−0.24***	0.38***	0.04	0.32***	-			
9. ARHR	0.06	−0.10	0.30***	−0.31***	0.44***	0.18*	0.12	0.28***	-		
10. RHW	0.12	0.02	0.38***	−0.04	0.29***	0.13	0.01	0.17*	0.50***	-	
11. CI (T1)	−0.01	0.07	0.24***	−0.48***	0.41***	0.21**	0.14*	0.33***	0.25**	0.08	-
12. CI (T2)	0.01	0.00	0.29***	−0.41***	0.43***	0.22**	0.13	0.32***	0.31***	0.23**	0.65***

Abbreviations: ARHR, analytical racial hoax reading; CI (T1), contact intention at T1; CI (T2), contact intention at T2; Credib, credibility; FC, fact-checking intention; PAR, propensity to engage in analytical reasoning; RHW, racial hoax re-writing; SelfTr, self-Transcendence; Share, intention to share the news; Ster., stereotypical belief.

*P < .05; **P < .01; ***P < .001.

that baseline contact intentions ($\beta = .57$, $P < .001$) and performances variables ($\beta = .21$, $P < .001$) were the only significant predictors of post-intervention contact intentions.

3.2 | Cluster analysis

Skewness and kurtosis values for the three dimensions of interest for cluster analysis fall within the range ± 2, thus indicating only slight deviations from a normal distribution that can be dealt with by Ward's hierarchical clustering method (1963).

Inspection of the dendrogram obtained from the replicated hierarchical clustering procedure through 5,000 artificial bootstrap samples preliminarily indicated a clear preference for a three-cluster solution. The evaluations of Dunn's index, mean Silhouette value, and connectivity index confirmed the stability of the three-cluster solution: both Dunn's index and the mean Silhouette value reached their maxima, and the connectivity index recorded its minimum, each for the three-cluster solution. The values of APN and ADM also were minimized at the three-cluster solution, confirming its validity. Therefore, based on that solution, the centroids of the three clusters were extracted and used as seed points to initialize a k-means procedure, definitively reclassifying the subjects into groups. The final group means are shown in Figure 3.

The results of the multivariate analysis of variance (Table 4) showed a significant effect due to the membership to the three different profiles of vulnerability to RM based on the three dimensions ($F(6, 406) = 115.03$, $P < .001$, Wilks' $\Lambda = 0.14$, $\eta^2 = 0.62$). The results of univariate ANOVAs show a significant effect of cluster membership on PAR ($F(2, 205) = 225.20$, $P < .001$, $\eta^2 = 0.69$), stereotypical beliefs ($F(2, 205) = 31.75$, $P < .001$, $\eta^2 = 0.24$) and also self-transcendence value ($F(2, 205) = 96.01$, $P < .001$, $\eta^2 = 0.48$).

Adolescents belonging to the first group ('Reflexive and not biased', 68 subjects, 32.7% of the total) show the highest levels of PAR, significantly higher than group two (1.77, $P < .001$) and group three (1.72, $P < .001$). The levels of stereotypical beliefs are very close to the mean of the total sample, lower than group three (0.86, $P < .001$) but significantly higher than group two (0.37, $P < .001$). The self-transcendence value of this group differs from that of group three, being significantly higher (1.55, $P < .001$), but not from that of the second group.

In contrast, group number two ('Impulsive and other-oriented', 89 subjects, 42.8% of the total) shows significantly lower mean values of PAR than group one, but not different from those found for group three. The level of
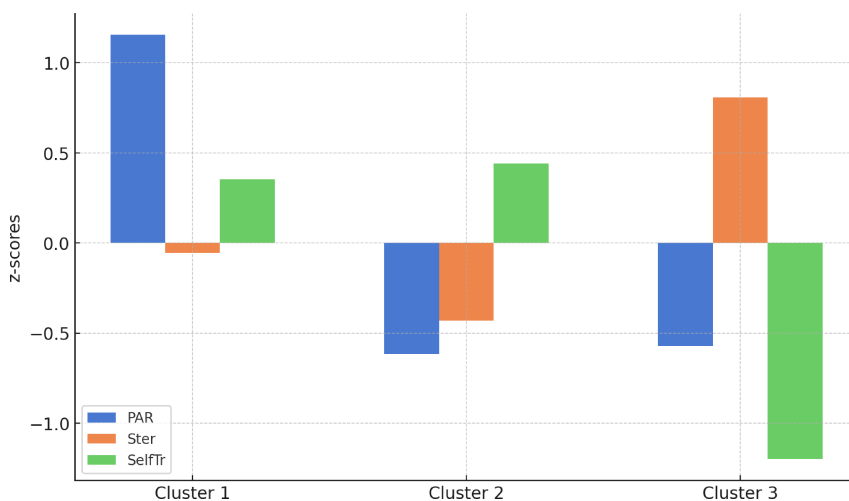

FIGURE 3 Plotted centroids of the three-cluster solution. PAR, propensity to engage in analytical reasoning; Ster, stereotypical belief; SelfTr, self-transcendence. Standardized centres are shown.

stereotypical beliefs of these adolescents, on the other hand, is the lowest among that of the other two groups, being lower than both group one and three (1.23, *P* < .001). As for self-transcendence, its value is not different from that of group one, but still significantly higher than that found in group three (1.64, *P* < .001).

Finally, adolescents belonging to group three ('Impulsive and biased', 51 subjects, 24.5% of the total) show a PAR not different from those in group two, but lower than those in group one. Stereotypical beliefs appear to be the highest for this group, differing significantly from the other two, and also the value of self-transcendence shows the lowest importance relative to the other two groups.

3.3 | Profile analysis

Considering the criterion variables considered, a significant multivariate effect among the groups is found (*F*(12, 262) = 7.71, *P* < .001; Wilks' Λ = 0.55, η^2 = 0.24). The three groups of adolescents (Figure 4) were not found to be different in their assessment of credibility of the news encountered (*F*(2, 205) = 2.29, *P* = .104, η^2 = 0.02). Regarding behavioural intention to share the news, univariate analysis of variance showed no significant result as

TABLE 4 Means and standard deviations of standardized values for each cluster and results of univariate analysis of variance.

	Cluster 1	Cluster 2	Cluster 3	*F*	η^2
PAR	1.15 (0.00)	−0.06 (1.02)	0.35 (0.77)	225.20***	0.69
Ster	−0.61 (0.63)	−0.43 (0.77)	0.44 (0.61)	31.75***	0.24
SelfTr	−0.57 (0.76)	0.81 (0.89)	−1.20 (0.82)	96.01***	0.48

Abbreviations: PAR, propensity to engage in analytical reasoning; Ster, stereotypical belief; SelfTr, self-transcendence.
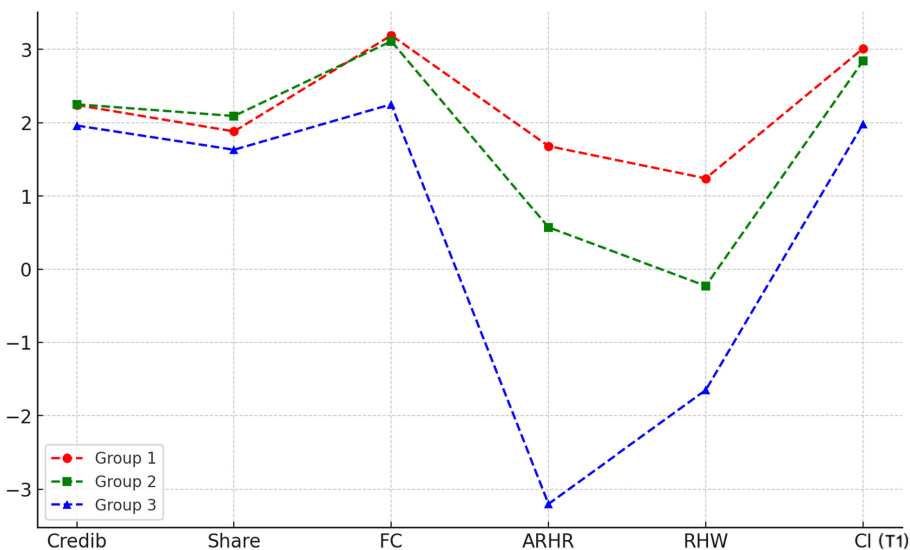***P < .001.

FIGURE 4 Profile plot showing the different responses to the considered criterion variables for the three groups. Credib, credibility; Share, intention to share the news; FC, fact-checking intention; ARHR, analytical racial hoax reading; RHW, racial hoax re-writing; CI (T1), contact intention at T1. Group 1, reflexive and not biased group; Group 2, impulsive and other-oriented group; Group 3, impulsive and biased group.

TABLE 5 Means and standard deviations of criterion variables for each group, with results of univariate analysis of variance.

	Group 1	Group 2	Group 3	F	η^2
Credib	2.24 (0.77)	2.25 (0.8)	1.96 (0.89)	2.29	0. 02
Share	1.88 (1.19)	2.09 (1.1)	1.63 (1.04)	2.80	0.03
FC	3.19 (0.90)	3.11 (0.95)	2.25 (1.18)	15.57***	0.13
ARHR	1.68 (3.02)	0.57 (4.14)	−3.20 (5.17)	17.12***	0.19
RHW	1.24 (2.66)	−0.23 (2.8)	−1.65 (2.87)	12.04***	0.14
CI (T1)	3.01 (0.81)	2.84 (0.79)	1.98 (0.97)	24.35***	0.19

Abbreviations: ARHR, analytical racial hoax reading; CI (T1), contact intention at T1; Credib, credibility; FC, fact-checking intention; Group 1, reflexive and not biased group; Group 2, impulsive and other-oriented group; Group 3, impulsive and biased group; RHW, racial hoax re-writing; Share, intention to share the news.
***$P < .001$.

well ($F(2, 205) = 2.80$, $P = .063$, $\eta^2 = 0. 03$), while significant univariate differences emerged between adolescent groups for fact-checking intention ($F(2, 205) = 15.57$, $P < 0.001$, $\eta^2 = 0.13$). Variables related to the socio-cognitive processes activated by the web app were significantly different between the adolescent groups at the univariate analysis of variance, both for analytical performance ($F(2, 150) = 17.12$, $P < .001$, $\eta^2 = 0.19$) and rewriting performance ($F(2, 154) = 12.04$, $P < .001$, $\eta^2 = 0.14$). Finally, contact intentions are significantly different between the three groups of adolescents as well: $F(2, 205) = 24.35$, $P < .001$, $\eta^2 = 0.19$ (Table 5).

Following the significant results of the analyses of variance, Tukey's honestly significant difference (HSD) post hoc tests were conducted to examine pairwise differences in the criterion variables between clusters. Regarding fact-checking intentions, a significant difference was found regarding adolescents belonging to the impulsive and biased group in relation to both reflexive and not biased (0.94, $P < .001$) and impulsive and other-oriented (0.86, $P < .001$), being the least intentional to perform fact-checking on the presented news. Analytical performance at news reading shows the same pattern, with the impulsive and biased group performing less well on average than reflexive and not biased (4.88, $P < .001$) and impulsive and other-oriented adolescents (3.77, $P < .001$). On rewriting following mediated contact, however, it is the reflexive and not biased group that performs significantly better (Group 1–Group 2:1.47, $P = .017$; Group 1–Group 3:2.89, $P < .001$). The difference between the performance at rewriting obtained on average by the impulsive and other-oriented group is also significantly better than that obtained by the impulsive and biased group (1.42, $P = .026$). Finally, the average contact intention scores also see the impulsive and biased group differ from the other two, being significantly lower than both (Group 1–Group 3:1.03, $P < .001$; Group 2–Group 3:0.86, $P < .001$).

3.4 | Path analysis

Prior to the path analysis, we assessed the distribution of the differences between paired observations of contact intention before and after the web app-based intervention for each of the three groups of adolescents. The results show a significant departure from the assumption of normality in the Shapiro–Wilk univariate test for the pre-post differences across the three clusters (Group 1:0.92, $P = .001$; Group 2:0.92, $P < .001$; Group 3:0.95, $P = .042$). Consequently, the Wilcoxon signed-rank test was employed as a non-parametric alternative to the paired-sample t-test. Although an increase in contact intentions was observed for each of the three groups, it was possible to conclude that contact intention scores after the socio-analytical intervention were significantly higher than before only for adolescents belonging to the impulsive and biased group (Group 1: $V = 312$, $P = .093$, two-tailed; Group 2: $V = 975$, $P = .231$, two-tailed; Group 3: $V = 328.5$, $P = .045$, two-tailed).

Mardia's test for multivariate normality (1970) indicated a significant departure from the assumption of multivariate normality in terms of both skewness and kurtosis (respectively, 79.29, $P < .01$; 2.70, $P = .007$). Consequently, path analysis was conducted using maximum likelihood estimation with robust (Huber–White) standard errors to accommodate this. Regarding missing data, Little's MCAR test (1988) yielded a non-significant result ($\chi^2 = 37.3$, $P = .113$), providing no basis to reject the hypothesis that the missing data mechanisms in the variables studied were missing completely at random. In line with this assumption (see Marcoulides & Schumacker, 2013), full information maximum likelihood (FIML) estimation was employed to address the missing data within the lavaan package (Muthén & Shedden, 1999; Schafer & Graham, 2002).

The plausibility of the relationships between variables was confirmed by the fit indices of the proposed model, following the established cut-off values proposed by Kline (2023):b $\chi^2 = 10.038$, df = 7, $P = .186$; CFI = .988; TLI = .976; RMSEA = .047 (90% CI [0.000, 0.113]), $P = .459$; SRMR = .038. Figure 5 shows the diagram with the standardized parameters.

The results of the path analysis highlight how the link from analytical racial hoaxes reading to racial hoaxes rewriting is positive and statistically significant ($\beta = 0.40$ [0.07], $P < .001$): this suggests that more analytical engagement with racial hoaxes is associated with greater capability in rewriting these hoaxes without racial manipulations. Compared to the reflexive and not biased group, belonging to the impulsive and other-oriented group is negatively associated with racial hoaxes rewriting ($\beta = -0.21$ [0.08], $P = .011$) but not with analytical performance, confirming that adolescents in the second group consistently obtained a lower performance in rewriting racial hoaxes after mediated contact. Conversely, belonging to the impulsive and biased group is negatively associated with analytical racial hoax reading ($\beta = -0.40$ [0.07], $P < .001$) and rewriting ($\beta = -0.23$ [0.09], $P = .01$), showing that adolescents in this third group show significantly worse performance in both analysing racial hoax manipulations and rewriting racial hoaxes after mediated contact. Furthermore, being part of the impulsive and biased group is negatively related to the level of pre-intervention contact intentions ($\beta = -0.43$ [0.06], $P < .001$), showing that participants in this third group were less likely to intend to make contact initially. The relationship between the baseline level of contact intentions and the same intentions measured at the second-time step is significant and consistent with what might be expected ($\beta = 0.63$ [0.06], $P < .001$). The link between racial hoax rewriting and the level of contact intentions measured after the intervention is also significant and positive ($\beta = 0.19$ [0.05], $P < .001$), highlighting that the act of rewriting racial hoaxes is potentially a predictor of the intention to subsequently engage in contact with members of the outgroup discriminated by the hoax.
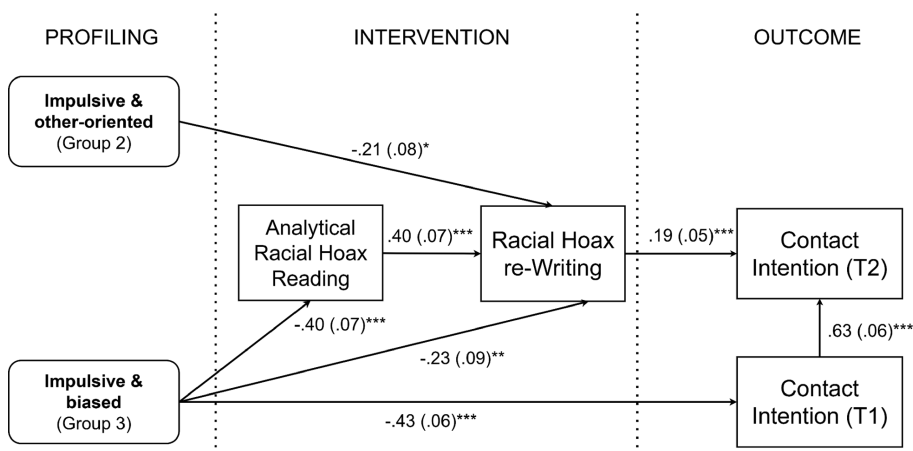

FIGURE 5 Path diagram with robust maximum-likelihood parameter estimates. The displayed estimates are the standardized coefficients with their standard errors (in parentheses). *$P < .05$; **$P < .01$; ***$P < .001$.

The indirect path from belonging to the impulsive and other-oriented group to post-intervention contact intentions via racial hoaxes rewriting is negative and significant (−0.04 [0.02], 95% CI [−0.077, −0.002]), indicating that, in comparison to those in the first group, adolescents in the impulsive and other-oriented group tend to have lower post-intervention contact intentions through their reduced performance in rewriting racial hoaxes. The indirect path from belonging to the impulsive and biased group to post-intervention contact intentions via analytical racial hoaxes reading and racial hoaxes rewriting is also negative and statistically significant (−0.08 [SE = 0.03], 95% CI [−0.124, −0.027]), showing that adolescents in this third group have overall lower post-intervention contact intentions (in comparison to those in the reflexive and not biased group) through their consistently poorer analytical and rewriting performance concerning racial hoaxes. All relationships in the model were controlled for the participant's gender and age, yet these variables did not exhibit significant associations with the other variables in the path analysis. Similarly, the level of contact intentions measured at T1 did not significantly impact the other variables in the model, except for the same variable at T2.

4 | DISCUSSION

The primary general objective of this contribution was to demonstrate how hybrid technology implemented within an educational community can enhance intergroup contact among adolescents' groups differently susceptible to potential anti-immigrant prejudice associated with racial hoaxes. Previous psycho-social research has already highlighted the effectiveness of mediated forms of contact in promoting for instance perspective taking and pro-immigrant attitudes (Banas et al., 2020; Goldman & Mutz, 2010). However, there was a lack of attention on how different adolescents can respond to this kind of intervention. Thus, to fill this gap the present study centres on identification of different adolescents' profiles resulting from the interplay between motivated reasoning (ethnic stereotypes and personal values; Kahan et al., 2017) and analytical reasoning (RQ1). The literature on both racial misinformation (D'Errico et al., 2022; Pennycook & Rand, 2019; Wright & Duong, 2021) and intergroup contact (Barni et al., 2020; Blanchar & Sparkman, 2020; Sagiv & Schwartz, 1995) has underscored the essential role of these motivational and cognitive factors in mitigating racial misinformation and promoting intergroup contact. The results of this study have highlighted their role in understanding how and why some adolescents are more vulnerable to racial misinformation, and how different prompts can be used in interventions aimed at fostering intergroup contact. Specifically, by adopting a person-centred approach, in response to RQ1 we identified groups of profiles that vary in their inclination to process information analytically (Pennycook & Rand, 2019) and employ differently distorted motivated cognitions (Kahan, 2013) when dealing with racial misinformation. Through the interaction of different dimensions (ie, PAR, self-transcendence values and stereotypes) we have identified three distinct types of young users: (1) Reflexive and not biased adolescents, comprising one-third of the sample, more adept at addressing racial misinformation in a reflexive and impartial manner. (2) Impulsive and other-oriented adolescents, constituting the majority of participants, who, despite positively valuing others, do not engage in an in-depth analysis of fake-news. (3) Impulsive and biased adolescents, making up a quarter of the sample, facing a higher risk when encountering such narratives due to their limited inclination for reflection, lesser focus on others and elevated levels of stereotyping. These profiles, in conjunction with both racial misinformation and intergroup contact intentions, can interact differently, defining mindsets that may influence behavioural reactions to such news (eg, fact-checking behaviours) and content processing during the intervention with the Rolling Minds web app, aimed at promoting socio-analytical reasoning and the adoption of alternative viewpoints. First, from the analysis of reactions of these different groups of adolescents to racial misinformation, it is noteworthy that all three groups, with no significant differences, may perceive racial misinformation as credible. Furthermore, all three groups exhibit a similar likelihood of sharing it. However, concerning fact-checking behaviour, only the impulsive and biased group is less likely to verify the credibility of sources and the news itself. These results suggest that there may be different reasons behind similar online behavioural and evaluative reactions related to racial hoaxes. For example, the reflexive and unbiased, and impulsive

and other-oriented groups may perceive the news as credible, but not necessarily true. Additionally, for these groups, sharing might be a form of reporting with prosocial intentions, defending the immigrants targeted by racial misinformation (D'Errico et al., 2022). On the other hand, for the impulsive and biased group the piece of news may align with their prejudices and be considered evidence of the validity of these negative evaluative attitudes (Wright & Duong, 2021). For this more vulnerable group, sharing intentions may be hostile and aimed at highlighting the perceived danger of immigrants without first verifying the reliability of the news content and sources. These findings emphasize the importance of exploring the presence of different mindsets in the pre-assessment phase of an intervention aimed at intergroup contact through countering racial misinformation. Regarding the intervention, and in particular the role of cognitive-motivational profiles in responding to the proposed interactive conversational intervention (RQ2), the results indicate that different adolescents' groups exhibit varying responses to the suggestions presented by the artificial peer educator (Sofia). Throughout the intervention phases, Sofia prompts users to engage in a series of socio-cognitive operations focused on media biases (Paul & Elder, 2006), and at the same time inducing mediated contact with the outgroup member (Banas et al., 2020), that is, Said. These operations aim to encourage reflection on the characteristics of racial misinformation and a reprocessing of their perspective on misleading racial news, thereby expanding the range of possible interpretations. Specifically, the present study reveals that adolescents in the reflexive and unbiased group, who initially show higher contact intentions, demonstrate a higher level of performance in both the reading and rewriting phases following mediated intergroup contact. For these adolescents, the intervention can support contact intentions, however, it is worth noting that this group exhibits a lower vulnerability mindset both in terms of susceptibility to the influence of racial misinformation and the risk of engaging in discriminatory behaviours. Regarding adolescents in the impulsive and other-oriented group, their levels of socio-analytical performance are lower compared to the previous group, which is more protected. Furthermore, an examination of the influence pattern of socio-analytical performance on contact intentions reveals that the rewriting phase is more challenging for them, resulting in less effective performance compared to those who are more reflective. For this group, the intervention is likely to primarily counter the effects of misinformation rather than promote contact intentions. This group appears to be protected from the risk of engaging in discriminatory behaviours but is more susceptible to becoming well-intentioned spreaders of racial misinformation on social media. It is likely that these adolescents, who are sensitive to civic and social issues, respond to this type of news with emotional activation that is not regulated by colder cognitive processes, making them less impartial and rational in information processing. In this case, the intervention could assist them in not only better recognizing racial hoaxes, but also articulating their prosocial orientation in a more reflective manner. Finally, concerning adolescents belonging to the impulsive and biased group, both contact intentions and socio-analytic performance are lower than those of all other groups. For them, the risk is twofold: being influenced by racial misinformation and increasing hostile perceptions towards immigrants. This result for the most at-risk group aligns with findings in the literature on misinformation (D'Errico et al., 2022; Pennycook & Rand, 2019; Wright & Duong, 2021) and intergroup contact (Barni et al., 2020; Blanchar & Sparkman, 2020; Sagiv & Schwartz, 1995). However, even though it is more challenging for them to respond to the modules proposed in the conversational app, the guided socio-analytical elaboration leads to the most pronounced improvement in intergroup contact intentions.

In summary, an examination of the responses to the intervention from different groups reveals the role of cognitive dimensions (PAR) and socio-motivational factors (values and stereotypes; Kahan, 2013), in identifying groups at risk of credulity during a pre-assessment phase. This phase could be useful for better tailoring the hybrid intervention, providing educators with insights about which aspects and modules to focus on particularly, based on the different difficulties present in the classes. For example, in the case of a majority of students with high self-transcendence and low inclination towards reflection, more time and space could be dedicated to analytical reflection on news, training students to recognize social influences online. On the other hand, students in the impulsive and biased group would benefit from a combination of perspective-taking interventions, like the mediated contact (Goldman & Mutz, 2010), that can also promote stereotypes recognition abilities and outgroup humanization, as well as analytical reflection on news. Furthermore, regardless of the difficulties, the results emphasize the malleability of students'

resources that can be developed through interaction with a conversational agent designed for educational purposes in real classroom contexts. Indeed, even for students at higher risk, hybrid technology (Gatti & Procentese, 2020) designed to engage young people in dialogue about the negative effects of technologically mediated interactions can be helpful in changing their approach to social issues (often polarizing) in a more critical and open manner.

This study, while having several positive aspects, also exhibits limitations due to its exploratory nature, primarily focused on testing a web app designed to induce media bias reflection and mediated intergroup contact. In this context, it could be useful for future studies to include a control group that interacts with different emerging profiles. Additionally, as previously described, this type of contact based on reading or watching the point of view of a fixed outgroup member can surely promote intergroup contact (Barni et al., 2020), and it provides a solid foundation for future studies: in particular, it has recently been highlighted that another kind of mediated contact, the so-called digital one (Amichai-Hamburger & McKenna, 2006), can also promote these positive effects on the users, thus future studies could benefit from exploring direct conversational interactions with outgroup members. Furthermore, future studies could certainly benefit from including, during the debriefing phase, a critical perspective on online learning environments (Nagle, 2018), which sometimes are unreal neutral spaces where interactions are induced and controlled, unlike social media, which can expose users to potentially aggressive and hateful attacks that must be recognized and distanced. In this phase, researchers and educators can draw on critical social media literacy, which points out that actual social media platforms are characterized by risky affordances (Rheingold, 2010) and a business model (Friesen & Lowe, 2012) where advertisers aim to attract and manipulate users' attention, exposing them to various social media risks. Nevertheless, in practical terms, this methodological approach (firstly, identifying risky profiles, and then intervening through conversational interactions in a hybrid digital context) could be significant in designing interventions to counter other types of misleading news targeting different minorities (eg, sexist or homophobic). Thus, future studies could explore how these findings might be applied in other settings, such as designing training for organizational contexts or targeting diverse samples, including young adults or the elderly.

AUTHOR CONTRIBUTIONS

Francesca D'Errico: Conceptualization; writing—original draft; methodology; funding acquisition. **Giuseppe Corbelli:** Data curation; formal analysis; software; writing—original draft. **Paolo Giovanni Cicirelli:** Data curation; methodology; investigation; writing—original draft. **Carmela Sportelli:** Writing—original draft; investigation. **Marinella Paciello:** Conceptualization; writing—original draft; methodology; supervision. All authors approved the final version of the article.

ACKNOWLEDGEMENTS

FUNDING INFORMATION

CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interests related to this publication.

DATA AVAILABILITY STATEMENT

The authors are willing to share their data, analytics methods and study materials with other researchers. The data supporting this study's findings are available from the corresponding author, P.G.C., upon request.

ETHICS STATEMENT

The study was conducted in accordance with the Declaration of Helsinki and aligned with the ethical code of the AIP. Furthermore, both the design and the procedures received approval from the ethics committee at the University, to which three of the authors are affiliated (reference code: ET-22-01).

INFORMED CONSENT

Informed consent was obtained from all individual participants included in the study.

ORCID

Francesca D'Errico https://orcid.org/0000-0002-8957-665X

Giuseppe Corbelli https://orcid.org/0000-0002-2864-3548

Paolo Giovanni Cicirelli https://orcid.org/0000-0001-5002-4834

Carmela Sportelli https://orcid.org/0009-0003-6573-3635

Marinella Paciello https://orcid.org/0000-0002-9023-0656

REFERENCES

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, *31*(2), 211–236.

Allport, G. W. (1954). The Nature of Prejudice (Addison–Wesley, 1954).

Amichai-Hamburger, Y., & McKenna, K. Y. (2006). The contact hypothesis reconsidered: Interacting via the internet. *Journal of Computer-Mediated Communication*, *11*(3), 825–843.

Athota, V. S., Kearney, S. P., & Cocodia, E. (2015). How self-transcendence via individualised moral foundations predict emotional and social enhancement. *Journal of Beliefs & Values*, *36*(3), 297–307.

Banas, J. A., Bessarabova, E., & Massey, Z. B. (2020). Meta-analysis on mediated contact and prejudice. *Human Communication Research*, *46*(2–3), 120–160.

Barni, D., Cavazza, N., Russo, S., Vieno, A., & Roccato, M. (2020). Intergroup contact and prejudice toward immigrants: A multinational, multilevel test of the moderating role of individual conservative values and cultural embeddedness. *International Journal of Intercultural Relations*, *75*, 106–117.

Birtel, M. D., Di Bernardo, G. A., Stathi, S., Crisp, R. J., Cadamuro, A., & Vezzali, L. (2019). Imagining contact reduces prejudice in preschool children. *Social Development*, *28*(4), 1054–1073.

Blanchar, J. C., & Sparkman, D. J. (2020). Individual differences in miserly thinking predict endorsement of racial/ethnic stereotypes. *Social Cognition*, *38*(5), 405–421.

Brock, G., Pihur, V., Datta, S., & Datta, S. (2008). *clValid*: An R package for cluster validation. *Journal of Statistical Software*, *25*, 1–22.

Cameron, L., Rutland, A., & Brown, R. (2007). Promoting children's positive intergroup attitudes towards stigmatized groups: Extended contact and multiple classification skills training. *International Journal of Behavioral Development*, *31*(5), 454–466.

Capanna, C., Vecchione, M., & Schwartz, S. H. (2005). La misura dei valori. Un contributo alla validazione del Portrait Values Questionnaire su un campione italiano. *Bollettino di Psicologia Applicata*, *246*, 29.

Cerase, A., & Santoro, C. (2018). From racial hoaxes to media hypes. Fake news' real consequences. In P. Vasterman (Ed.), *From Media Hype to Twitter Storm: News Explosions and Their Impact on Issues, Crises and Public Opinion* (pp. 333–354). Amsterdam, The Netherlands: AUP-Amsterdam University Press.

Corbelli, G. (2024). reflectR: Automatic Scoring of the Cognitive Reflection Test. R package version 2.1.3. https://CRAN.R-project.org/package=reflectR

Corenblum, B., & Stephan, W. G. (2001). White fears and native apprehensions: An integrated threat theory approach to intergroup attitudes. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, *33*(4), 251.

Cuddy, A. J., Fiske, S. T., Kwan, V. S., Glick, P., Demoulin, S., Leyens, J. P., ... Ziegler, R. (2009). Stereotype content model across cultures: Towards universal similarities and some differences. *British Journal of Social Psychology*, *48*(1), 1–33.

D'Errico, F., Cicirelli, P. G., Corbelli, G., & Paciello, M. (2023). Addressing racial misinformation at school: A psycho-social intervention aimed at reducing ethnic moral disengagement in adolescents. *Social Psychology of Education*, *27*, 611–630. https://doi.org/10.1007/s11218-023-09777-z

D'Errico, F., Corbelli, G., Papapicco, C., & Paciello, M. (2022). How personal values count in misleading news sharing with moral content. *Behavioral Sciences*, *12*(9), 302.

Edgerly, S., Mourão, R. R., Thorson, E., & Tham, S. M. (2020). When do audiences verify? How perceptions about message and source influence audience verification of news headlines. *Journalism & Mass Communication Quarterly*, *97*(1), 52–71.

Faragó, L., Orosz, G., Paskuj, B., & Krekó, P. (2024). Analytic adolescents prevail over fake news – A large-scale preregistered study. *Personality and Individual Differences*, 229, 112747. https://doi.org/10.1016/j.paid.2024.112747

Figenschou, T. U., & Thorbjørnsrud, K. (2015). Faces of an invisible population: Human interest framing of irregular immigration news in the United States, France, and Norway. *American Behavioral Scientist*, 59(7), 783–801.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, 19(4), 25–42.

Friesen, N., & Lowe, S. (2012). The questionable promise of social media for education: Connective learning and the commercial imperative. *Journal of Computer Assisted Learning*, 28(3), 183–194.

Gatti, F., & Procentese, F. (2020). Open neighborhoods, sense of community, and Instagram use: Disentangling modern local community experience through a multilevel path analysis with a multiple informant approach. *TPM: Testing, Psychometrics, Methodology in Applied Psychology*, 27(3), 313–329.

Gatti, F., Procentese, F., & Schouten, A. P. (2023). People-nearby applications use and local community experiences: Disentangling their interplay through a multilevel, multiple informant approach. *Media Psychology*, 26(3), 278–305.

Goldman, S. K., & Mutz, D. C. (2010). Effects of mass media on prejudice. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping and discrimination* (pp. 167–181). New York: Psychology Press.

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology*, 101(2), 366.

Greene, C. M., & Murphy, G. (2023). Debriefing works: Successful retraction of misinformation following a fake news study. *PLoS One*, 18(1), e0280295.

Hadarics, M., & Kende, A. (2018). The dimensions of generalized prejudice within the dual-process model: The mediating role of moral foundations. *Current Psychology*, 37, 731–739.

Harrell, F. E., Jr. (2024). *Hmisc*: Harrell Miscellaneous. R package version 5.1-2. https://CRAN.R-project.org/package=Hmisc

Hennig, C. (2023). fpc: Flexible Procedures for Clustering. R package version 2.2-11. https://CRAN.R-project.org/package=fpc

Herrero-Diz, P., Conde-Jiménez, J., & Reyes-de-Cózar, S. (2021). Spanish adolescents and fake news: Level of awareness and credibility of information (los adolescentes españoles frente a las fake news: Nivel de conciencia y credibilidad de la información). *Culture and Education*, 33(1), 1–27.

Howard, M. C., & Hoffman, M. E. (2018). Variable-centered, person-centered, and person-specific approaches: Where theory meets the method. *Organizational Research Methods*, 21(4), 846–876.

Hurtado, S. (2005). The next generation of diversity and intergroup relations research. *Journal of Social Issues*, 61(3), 595–610.

Kahan, D. M. (2013). Ideology, motivated reasoning, and cognitive reflection. *Judgment and Decision Making*, 8(4), 407–424.

Kahan, D. M., Peters, E., Dawson, E. C., & Slovic, P. (2017). Motivated numeracy and enlightened self-government. *Behavioural Public Policy*, 1(1), 54–86.

Kassambara, A., & Mundt, F. (2020). factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R package version 1.0.7. https://CRAN.R-project.org/package=factoextra

Kline, R. B. (2023). *Principles and practice of structural equation modeling*. New York: Guilford Publications.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480.

Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83(404), 1198–1202.

Lutzke, L., Drummond, C., Slovic, P., & Árvai, J. (2019). Priming critical thinking: Simple interventions limit the influence of fake news about climate change on Facebook. *Global Environmental Change*, 58, 101964.

MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, 39(1), 99–128.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2023). cluster: Cluster Analysis Basics and Extensions. R package version 2.1.6. https://CRAN.R-project.org/package=cluster

Magnusson, D., & Stattin, H. (1996). *Person-context interaction theories*. Stockholm: Department of Psychology, University of Stockholm.

Maio, G. R., & Olson, J. M. (1998). Values as truisms: Evidence and implications. *Journal of Personality and Social Psychology*, 74(2), 294.

Marcoulides, G. A., & Schumacker, R. E. (2013). *Advanced structural equation modeling: Issues and techniques*. New York: Psychology Press.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57(3), 519–530.

Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*, 55(2), 463–469.

Nagle, J. (2018). Twitter, cyber-violence, and the need for a critical social media literacy in teacher education: A review of the literature. *Teaching and Teacher Education*, 76, 86–94.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175–220.

Nir, L. (2011). Motivated reasoning and public opinion perception. *Public Opinion Quarterly*, 75(3), 504–532.

Papapicco, C., Lamanna, I., & D'Errico, F. (2022). Adolescents' vulnerability to fake news and to racial hoaxes: A qualitative analysis on Italian sample. *Multimodal Technologies and Interaction*, 6(3), 20.

Paravati, E., Fitzgerald, K., Green, M. C., McAllister, C., & Moore, M. M. (2022). Narratives to increase prosociality toward refugees. *International Journal of Communication*, 16, 22.

Paul, R., & Elder, L. (2006). *How to detect media bias and propaganda*. Dillon Beach, CA: Foundation for Critical Thinking.

Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, 188, 39–50.

Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, 25(5), 388–402.

Pereira, A., Harris, E., & Van Bavel, J. J. (2023). Identity concerns drive belief: The impact of partisan identity on the belief and dissemination of true and false news. *Group Processes & Intergroup Relations*, 26(1), 24–47.

Pettigrew, T. F. (1982). *Prejudice*. Cambridge, MA: Harvard University Press.

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, 90(5), 751–783.

R Core Team. (2021). *R: A language and environment for statistical computing [Computer software]*. R Foundation for Statistical Computing. https://www.R-project.org/

Rahn, W. M. (1993). The role of partisan stereotypes in information processing about political candidates. *American Journal of Political Science*, 37, 472–496.

Revelle, W. (2024). *Psych: Procedures for psychological, psychometric, and personality research*. Evanston, Illinois: Northwestern University. R package version 2.4.3. https://CRAN.R-project.org/package=psych

Rheingold, H. (2010). Attention, and other 21st-century social media literacies.

Rosenblatt, A., Greenberg, J., Solomon, S., Pyszczynski, T., & Lyon, D. (1989). Evidence for terror management theory: I. The effects of mortality salience on reactions to those who violate or uphold cultural values. *Journal of Personality and Social Psychology*, 57(4), 681.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48, 1–36.

Sagiv, L., & Schwartz, S. H. (1995). Value priorities and readiness for out-group social contact. *Journal of Personality and Social Psychology*, 69(3), 437.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7, 147–177.

Souchon, N., Maio, G. R., Hanel, P. H., & Bardin, B. (2017). Does spontaneous favorability to power (vs. universalism) values predict spontaneous prejudice and discrimination? *Journal of Personality*, 85(5), 658–674.

Stephan, C. W., Stephan, W. C., Demitrakis, K. M., Yamada, A. M., & Clason, D. L. (2000). Women's attitudes toward men an integrated threat theory approach. *Psychology of Women Quarterly*, 24(1), 63–73.

Stephan, W. G., Boniecki, K. A., Ybarra, O., Bettencourt, A., Ervin, K. S., Jackson, L. A., ... Renfro, C. L. (2002). The role of threats in the racial attitudes of Blacks and Whites. *Personality and Social Psychology Bulletin*, 28(9), 1242–1254.

Szebeni, Z., Jasinskaja-Lahti, I., Lönnqvist, J. E., & Szabó, Z. P. (2023). The price of (dis) trust–profiling believers of (dis) information in the Hungarian context. *Social Influence*, 18(1), 2279662.

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, 50(3), 755–769.

Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making*, 11(1), 99–113.

Traberg, C. S., Roozenbeek, J., & van der Linden, S. (2022). Psychological inoculation against misinformation: Current evidence and future directions. *The Annals of the American Academy of Political and Social Science*, 700(1), 136–151.

Tropp, L. R., & Pettigrew, T. F. (2005). Differential relationships between intergroup contact and affective and cognitive dimensions of prejudice. *Personality and Social Psychology Bulletin*, 31(8), 1145–1158.

Vezzali, L., Hewstone, M., Capozza, D., Giovannini, D., & Wölfer, R. (2017). Improving intergroup relations with extended and vicarious forms of indirect contact. In *European review of social psychology* (Vol. 25, pp. 314–389). New York: Routledge.

Visintin, E. P., Voci, A., Pagotto, L., & Hewstone, M. (2017). Direct, extended, and mass-mediated contact with immigrants in Italy: Their associations with emotions, prejudice, and humanity perceptions. *Journal of Applied Social Psychology*, 47(4), 175–194.

Ward, J. H., Jr. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301), 236–244.

Wickham, H., Miller, E., & Smith, D. (2023). haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files. R package version 2.5.4. https://CRAN.R-project.org/package=haven

Wright, C., Brinklow-Vaughn, R., Johannes, K., & Rodriguez, F. (2021). Media portrayals of immigration and refugees in hard and fake news and their impact on consumer attitudes. *Howard Journal of Communications*, 32(4), 331–351.

Wright, C. L., & Duong, H. (2021). COVID-19 fake news and attitudes toward Asian Americans. *Journal of Media Research*, 14(1), 5–29.

Yilmaz, O., Karadöller, D. Z., & Sofuoglu, G. (2016). Analytic thinking, religion, and prejudice: An experimental test of the dual-process model of mind. *The International Journal for the Psychology of Religion*, 26(4), 360–369.

Zhang, Y., Zhou, M., Shao, Y., & Zhang, M. Y. (2020). mvnormalTest: Powerful tests for multivariate normality. R package version, 1(0), 390.