



Strategies to combat misinformation: Enduring effects of a 15-minute online intervention on critical-thinking adolescents

Gábor Orosz^{a,1}, Laura Faragó^{b,*,1}, Benedek Paskuj^c, Péter Krekó^{b,d}

^a ULR 7369 - URePSSS - Unité de Recherche Pluridisciplinaire Sport Santé Société, Sherpas, Univ. Lille, Univ. Artois, Univ. Littoral Côte d'Opale, Liévin, France

^b Institute of Psychology, ELTE Eötvös Loránd University, Budapest, Hungary

^c University College London, Department of Psychology, United Kingdom

^d Political Capital Institute, Budapest, Hungary

ARTICLE INFO

Handling editor: Nicolae Nistor

Keywords:

Adolescents
Wise intervention
Prosocial purpose
Family values
Fake news intervention
Counter-misinformation intervention

ABSTRACT

How is it possible to socialize adolescents to become more vigilant in spotting fake news? In the present pre-registered, randomized controlled trial ($N = 1476$), we aimed to implement a scalable, online counter-misinformation intervention by promoting family-based prosocial values and putting them in an expert role to build resistance against misinformation among adolescents. In this intervention, participants were endowed with an expert role and requested to write a letter to their digitally less experienced relatives elucidating six strategies to identify misinformation. We found immediate effects of the intervention ($d = 0.17$), but these effects disappeared after four weeks. However, those high school students who followed the instructions ($N = 791$) and had a higher need for cognition demonstrated a substantial benefit in correctly spotting fake news four weeks after the intervention compared to the control group ($d_{+1 \text{ SD need for cognition}} = 0.28$, $d_{+2 \text{ SD need for cognition}} = 0.51$). The present work demonstrates the power of using classic social psychological components, such as a digital mindset, expertise role, and prosocial strategies, to achieve long-term behavioral change among certain adolescents. Our approach might complement prior nudge and inoculation interventions in the fight against misinformation in this age group.

1. Introduction

Misinformation² has emerged as a cardinal social and political problem of the 21st century: it can distort democratic processes by hindering voters' informed decision-making about politics and social issues (Allcott & Gentzkow, 2017), resulting in alienation and cynicism (Balmas, 2014), and reducing collaboration in crises that require both individual- and group-level cooperation, like the climate change or the COVID pandemic (Allington, Duffy, Wessely, Dhavan, & Rubin, 2021; Lewandowsky, Stritzke, Freund, Oberauer, & Krueger, 2013). Hence, social scientists, in particular psychologists, have shown great interest in

devising interventions that arm people against fake news by using accuracy nudging techniques (see e.g., Pennycook, Epstein, et al., 2021; Pennycook, McPhetres, Zhang, Lu, & Rand, 2020; Pennycook & Rand, 2021), applying inoculation or building digital literacy skills (see e.g., Guess et al., 2020; Maertens, Anseel, & van der Linden, 2020; Maertens, Roozenbeek, Basol, & van der Linden, 2021; Roozenbeek, Van der Linden, S., & Nygren, 2020), or using wise interventions³ exploiting prosocial motivations (Orosz, Paskuj, Faragó, & Krekó, 2023). Psychological fake news interventions have promising results in diminishing the perceived accuracy of fake information, and some of them even have lasting effects, e.g., persisting for weeks or months (see e.g., Guess et al.,

* Corresponding author. Institute of Psychology, ELTE Eötvös Loránd University, Izabella street 46., 1064, Budapest, Hungary.

E-mail address: farago.laura@ppk.elte.hu (L. Faragó).

¹ Both authors share the first authorship of this manuscript.

² We used the terminology of Pennycook and Rand (2021) to distinguish among the concepts of fake news, misinformation, and disinformation. Nevertheless, our intention is not to suggest that online information is exclusively true or false; interpreting news content in terms of its accuracy isn't always straightforward. Nonetheless, in this study, our primary objective was to select unequivocally fake and real news content, distinctively belonging to one category or the other. Hence, we deemed this definition suitable for our research objectives.

³ The term 'wise intervention' denotes a particular form of intervention aimed at altering individuals' subjective meaning-making processes, as Walton and Wilson (2018) described. Not all interventions fall under this category. For instance, accuracy nudges are considered interventions but do not impact the aforementioned meaning-making processes; therefore, they cannot be classified as wise interventions.

<https://doi.org/10.1016/j.chb.2024.108338>

Received 5 February 2024; Received in revised form 6 June 2024; Accepted 8 June 2024

Available online 9 June 2024

0747-5632/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

2020; Maertens et al., 2020, 2021; Zerback, Töpfl, & Knöpfle, 2020). However, none of these prior intervention attempts focused on durable effects among adolescents. The present work aims to fill this gap by providing further empirical support for the efficiency of interventions that aim to empower individuals by putting them in an “expert” role and exploiting their prosocial motivations.

Though fake news interventions have been tested in various contexts (Pennycook & Rand, 2021; Roozenbeek, Van der Linden, & Nygren, 2020), the age group of adolescents is largely understudied. The news consumption patterns of adolescents indicate an increased vulnerability to fake news (Baptista, Gradim, & Correia, 2022; Marchi, 2012; Newman, Fletcher, Kalogeropoulos, & Nielsen, 2019) with a corresponding limited aptitude in the identification of misinformation (Common Sense Media, 2017; Papapicco, Lamanna, & D’Errico, 2022). Even though they are vulnerable, very few critical thinking and media literacy interventions have been published in academia about adolescents (Delaney, Bacon, & Matson, 2022; Smith & de los Santos, 2022), and only a limited number of studies have exclusively concentrated on addressing their susceptibility to misinformation (see e.g., Barzilai et al., 2023; Caroti, Adam-Troian, Bagneux, & Theraud, 2023, preprint). Furthermore, adolescents usually have lower motivation and engagement in social-psychological interventions than other age groups (see e.g., Yeager, Dahl, & Dweck, 2018), making it harder to influence their behavior. Therefore, we aim to fill this gap in the misinformation literature and introduce a new line of intervention targeting adolescents: a prosocial wise fake news intervention, and test its long-term effectiveness in decreasing the perceived accuracy of misinformation.

1.1. Adolescents’ vulnerability to misinformation

Multiple studies demonstrated that adolescents are vulnerable to misinformation. Even though they are soon-to-be voters, many teenagers are considered “news avoiders” (Shehata, 2016), and they often lack intrinsic motivation to read the news (Baptista et al., 2022; Tamboer, Kleemans, & Daalmans, 2022). Adolescents are less interested in news consumption through traditional news outlets compared to previous generations, and they rather turn to social media sites as their primary information source (Marchi, 2012; Newman et al., 2019), which can increase their susceptibility to misinformation (Baptista et al., 2022; Benaissa Pedriza, 2021). Due to their age, their critical thinking is not fully developed (Papapicco et al., 2022) despite being a crucial skill to recognize fake news efficiently (Faragó, Krekó, & Orosz, 2023; Pennycook & Rand, 2019). Though adolescents know that some news was created to manipulate readers, and they are confident that they can identify false content effectively (Corbu, Oprea, & Frunzaru, 2022; Herrero-Diz, Conde-Jiménez, & Reyes de Cózar, 2021; Papapicco et al., 2022), in practice, they are typically less successful at spotting misinformation (Common Sense Media, 2017; Papapicco et al., 2022). Susceptibility to misinformation among adolescents is related to a less rigorous treatment of the information (Baptista et al., 2022), a lack of sufficient concern regarding the accuracy of the news, being unmotivated to use time-consuming evaluation strategies to spot fake news (Paul, Macedo-Rouet, Rouet, & Stadler, 2017), and a lack of media literacy education (Kahne & Bowyer, 2017). For instance, when evaluating the credibility of information, adolescents rarely pay attention to fundamental criteria such as the author or source reliability (Herrero-Diz et al., 2021; Tamboer, Kleemans, & Daalmans, 2022; Tamboer, Kleemans, Molenaar, & Bosse, 2022; Wineburg, Breakstone, McGrew, & Ortega, 2016). Therefore, even though they were “born” to the social media-dominated world, they can have shortcomings in their digital literacy and evaluation skills required to differentiate real from fake content or identify information manipulation. Adolescents’ news consumption habits indicate a high susceptibility to misinformation. While vulnerability to misinformation is not limited to this age group, and people in higher age groups can also have problems in identifying fake news and correcting it due to cognitive factors and their less experience

in social media (Brashier & Schacter, 2020), the above-mentioned age-specific factors draw a special kind of vulnerability for adolescents. This problem requires scalable online interventions with long-lasting effects aiming to increase critical thinking and digital media literacy skills to arm them against fake news.

1.2. Critical thinking, media literacy, and counter-misinformation interventions among adolescents

Regarding critical thinking and media literacy interventions among adolescents, the evidence is very scarce, though there are some promising avenues in the literature. For instance, a qualitative study tested the effect of reading, discussing, and critiquing a series of historical texts to secondary school students during a reading intervention class to empower youth to successfully evaluate sources, challenge fake news, and make informed decisions (Delaney et al., 2022). Nevertheless, the authors only used thematic analyses to evaluate the effectiveness of this intervention and neither applied a control group nor compared pre-and post-test results of fake news evaluation accuracy ratings or cognitive reflection to quantitatively test its efficiency. In another study, the authors applied a breaking news simulation to improve adolescents’ digital literacy skills (Smith & de los Santos, 2022). In this intervention, news literacy awareness and news appreciation measures were addressed before, immediately after, and two weeks after taking part in the breaking news simulation, and the intervention group watched a video about a journalist explaining news skills and standards. Even though the simulation in itself did not work, the students belonging to the intervention group had a better concept of news literacy and reported more understanding of the value of news sourcing (Smith & de los Santos, 2022). Nevertheless, this intervention did not test the change in adolescents’ susceptibility to fake information. None of these attempts were randomized controlled trials that could assess the effectiveness of a counter-misinformation program in the field. However, some recent endeavors have employed randomized controlled trials and implemented either an inoculation-based game, namely “Misinformation is Contagious” developed by Barzilai et al. (2023), targeting 7th and 8th-grade students, or examined the efficacy of a critical thinking education intervention among secondary school students, as explored by Caroti et al. (2023, preprint). The study conducted by Barzilai et al. (2023) revealed that engagement with the gameplay led to enhanced accuracy and sharing discernment among middle school students when compared to the active control group. On the other hand, Caroti et al. (2023, preprint) demonstrated that their intervention resulted in an immediate and sustained reduction in conspiracy beliefs among participants post-intervention and even one month after. These interventions effectively cultivated students’ evaluation strategies and equipped them with the ability to identify unreliable information and, therefore, *how* to spot misinformation. However, neither of these interventions tested the long-term change in fake news evaluation accuracy specifically (the only exception is the study of Caroti et al. (2023, preprint), but they investigated conspiracy and paranormal beliefs). We aim to supplement this line of research with our wise intervention with its strong focus on prosocial values, and not only teach adolescents *how*, but also *why* fake news detection is important. Furthermore, our objective is to assess the intervention’s efficacy in decreasing the perceived accuracy of fake news one month after its implementation.

1.3. Implementing a wise prosocial fake news intervention attempt to adolescents

Wise interventions (Walton, 2014) represent a specific type of intervention, and the objective of their usage is to change a particular aspect of how individuals think or feel in their everyday lives, ultimately promoting their well-being. Compared to previous interventions, their uniqueness lies in their capacity for psychological precision, their brief nature, and their common purpose of modifying self-reinforcing

mechanisms that evolve gradually, thereby long-lastingly enhancing individuals' prospects across diverse situations (Walton, 2014). The success of wise interventions lies in changing the subjective meaning-making processes of individuals (how people make sense of themselves and social situations, Walton & Wilson, 2018). Therefore, wise intervention can be applied to alter specific underlying psychological mechanisms that either contribute to societal issues or hinder individuals from thriving (Walton, 2014; Walton & Wilson, 2018).

Wise psychological interventions (e.g., Reeves et al., 2021; Yeager et al., 2014) use prosocial values for changing meaning-making processes, which can inspire people to integrate personal (e.g., spotting fake news for myself) and prosocial goals (e.g., spotting fake news for helping another person). Research indicates that individuals motivated by prosocial concerns exhibit a heightened commitment to exert greater effort, prioritize safety measures, and collaborate more with others (e.g., Grant, 2007; Grant and Hofmann, 2011; Grant and Shandell, 2022; Paunesku et al., 2015; Reeves et al., 2021; Yeager et al., 2014). For example, in educational settings, students exhibited improved performance in tedious and monotonous tasks when they were driven by prosocial incentives, leading to improved academic performance in subsequent months (Paunesku et al., 2015; Reeves et al., 2021; Yeager et al., 2014). Prosocial wise interventions can be successfully applied in various domains, therefore, we presume that they can be efficiently exploited to help spot fake news among adolescents as well. In the Hungarian socio-cultural context, family and family ties have a central value (Csíste, 2009; Hankiss, 1990; Pátkainé Bende, 2022). Therefore, we expected that adapting a prosocial intervention by highlighting the importance of helping family members can encourage students effectively.

A new promising area in the misinformation literature is the wise prosocial fake news intervention that aims to encourage the younger generations to help their digitally less experienced family members, and this promotion of prosocial values thereby decreases young adults' susceptibility to misinformation (see also Orosz et al., 2023). In this intervention, young adults were put in an expert role in the digital world and were asked to write a letter to their digitally less familiar elder relatives, explaining six strategies to help them recognize fake news. In the letter-writing task, a "saying-is-believing" technique was applied: if people give advice, but their behavior is not aligned with their advice, cognitive dissonance arises, which motivates the advice-giver to behave following the advice, resulting in an indirect persuasion (Aronson, 1999; Aronson, Fried, & Stone, 1991; Higgins & Rholes, 1978; LaCrosse, Canning, Bowman, Murphy, & Logel, 2020; Stephens, Hamedani, & Destin, 2014; Yeager & Walton, 2011), and more vigilance in news consumption. By the advice-giving and promoting pro-social values, the intervention indeed reduced young adults' vulnerability to misinformation: compared to the active control group, the intervention group showed an immediate effect ($d = 0.32$) that persisted until the follow-up four weeks later ($d = 0.22$) on their fake news evaluation accuracy ratings (Orosz et al., 2023).

This randomized controlled study is among the first psychological fake news interventions exploiting prosocial values and could complement prior intervention approaches in the fight against misinformation (Orosz et al., 2023). As young adults were treated as experts rather than incompetent participants and were asked to help their relatives, they could fulfill their need for high status. These preconditions are necessary to increase adolescents' motivation and engagement in intervention studies (Yeager et al., 2018). Therefore, we think this prosocial fake news intervention could be successfully replicated among high school students, as it considers young people's needs and important motivations.

1.4. Putting the prosocial intervention into cultural and socio-political context

Hungary presents an opportune context for investigating interventions to combat misinformation. The mainstream media in

Hungary has been plagued by systematic pro-government disinformation campaigns (Bozóki & Hegedűs, 2018; Demeter, 2018; Guriev & Treisman, 2022; Krekó, 2022; Krekó & Enyedi, 2018), leading to the categorization of Hungary as an "informational autocracy" (Krekó, 2022). This political influence has eroded free and independent media, earning Hungary the "partly free" classification since 2019 (Freedom House, 2019). Over the years, Hungary has significantly dropped in the global rankings of media freedom and media pluralism sliding from the 25th position in 2009 to the 85th position in 2022 (Polyák, Urbán, & Szávai, 2022; Reporters Without Borders, 2009, 2022). An illustrative example of the pro-government media influence occurred in early 2023, when an analysis of the Breaking News section of the pro-government news portal, Origo, revealed striking findings (Political Capital, 2023). Thematic analysis of 727 articles published during the first year of the Russian-Ukraine war indicated a biased portrayal. The headlines depicted Ukraine and the United States as aggressors, portrayed Zelensky in a negative light, and depicted Putin as competent and striving for de-escalation. These headlines reflected the language employed by Soviet "peace movements" during the Cold War, employing oversimplifying rhetoric. Furthermore, they disseminated widely debunked disinformation propagated by the Russian propaganda machinery, which has a well-documented impact on public attitudes (Political Capital, 2023). Therefore, investigating susceptibility to fake news is of utmost importance in this informational autocracy, especially among the soon-to-be-voters age group: adolescents.

1.5. Current research and hypotheses

In the current study, we aimed to utilize these prosocial values rooted in familial bonds to increase the youth's attentiveness to misinformation. The present study is a replication of a previously successful prosocial counter-misinformation intervention (Orosz et al., 2023) using a multi-site, large-scale Hungarian high school sample. Our goal was to examine the immediate and long-term effects of this intervention on spotting misinformation and exploring potential socioeconomic and psychological moderators. In the present manuscript, we only report results related to the main hypotheses regarding fake news evaluation accuracy ratings (H1), and the potential moderating factors. In addition, we will examine the intent-to-treat effects including every participant who was assigned to one of the two conditions. We were interested in whether our intervention could effectively reduce fake news evaluation accuracy ratings in ideal circumstances, therefore, we examined the treatment-of-the-treated effects among those adolescents who took the intervention sufficiently seriously.

According to our preregistered main hypotheses,⁴ we supposed that participants of the treatment group (compared to the control) would evaluate fake news as less accurate (H1a) both immediately after the intervention and one month after (H1b). Furthermore, based on the preregistration, we aimed to explore the moderating effects of the relevant individual difference variables (analytic thinking, bullshit receptivity, digital literacy, need for cognition, socioeconomic status, and ethnic minority status) on immediate and long-term fake news evaluation accuracy.⁵

⁴ The present work is part of a larger research project in which we examine the avenue of prosocial wise interventions to make people willing to spot misinformation in the long run. The present manuscript focuses only on the first hypothesis (H1), in which we tested the effectiveness of the intervention on fake news evaluation accuracy. We reported the results of the other preregistered hypotheses as important individual factors making adolescents less susceptible to fake information in another manuscript (under review). See the anonymized preregistration here: https://osf.io/8tgk6/?view_only=6d28c169133f4d6bb421631b90e9441b.

⁵ For more details about these additional variables, see the Supplemental Material.

Based on the preregistration, we ran both the intent-to-treat (ITT) and the treatment-of-the-treated (ToT) analyses. The main difference between the two is that with the ITT we can examine whether the intervention is effective in the field under various potentially contaminating circumstances (e.g., implementation errors, missing and inadequate responses, etc., Shaya & Gu, 2007). However, with the ToT analysis, we can verify whether under ideal circumstances the intervention has the potential to have an impact on the participants (e.g., examining the effects among those participants who followed the instructions and elaborated on the task).

Therefore, the first analysis was the *intent-to-treat analysis* including everybody who provided data about their fake news evaluation accuracy. Unfortunately, we could not analyze the responses of those students who dropped out before the outcome measures, as we do not have relevant dependent variables. Based on the pretests, we prepared for the situation that a significant portion of the adolescent, high school student respondents will not take the task seriously. Multiple sources of resistance might occur in the classroom in such data gathering situation: there might be resistance towards news reading, students can be reluctant to respond to a survey, or they might not have an ideal relationship with the professor who provided the link and invited them to participate in the survey. As we expected these difficulties, based on the preregistered *treatment-of-the-treated analysis*, we analyzed the data of those students who followed the instruction of the intervention/control materials to be able to see whether the activated psychological mechanisms can lead to a long-term effect in spotting fake news in this under-investigated age-group in this digitally vulnerable country.

2. Methods

2.1. Participants

Based on the preregistration, our goal was to recruit at least 787 participants based on a-priori power analysis (with $d = 0.20$,⁶ alpha error prob = 0.05 with 80% power, and a number of predictors = 1, based on Guess et al., 2020) to be able to test our hypothesis about fake news evaluation accuracy. Altogether, 1582 high school students from 25 institutions in Budapest and the countryside started the questionnaire. Among them, 1476 students were randomly assigned to either the treatment ($n = 735$) or the control condition ($n = 741$), and 106 students withdrew from the completion before getting the intervention or control materials (for the consort diagram, see Fig. 1). Participants' ages ranged from 14 to 21 years ($M = 16.37$, $SD = 1.11$), 52.23% of them were girls, 45.05% were boys, and 2.71% indicated other. Regarding the type of residence, 35.43% lived in the capital city, 21.13% resided in larger county towns, 22.90% lived in smaller towns, and 20.53% resided in villages. Considering ethnic and socioeconomic status, 8.67% of students identified as an ethnic minority; 24.93% of students were first-generation, and 75.07% reported that either their mother or their father has a college or university degree.

During the follow-up, there was significant attrition because some students failed to provide their proper code (generated in the first data collection), making it impossible to match their follow-up responses with their first-round data. Furthermore, a new wave of the COVID pandemic broke out at the time of the follow-up study, so many students were absent from school. These obstacles led to the attrition of 43.6% of students in the follow-up data ($N = 832$). Of the remaining students ($M = 16.32$, $SD = 1.07$), 54.45% were girls, 42.66% were boys, and 2.88% indicated other. Regarding the place of settlement, 35.46% lived in Budapest, 19.35% resided in larger county towns, 24.03% lived in towns, and 21.15% resided in villages. Considering ethnic and socioeconomic status, 5.52% of students identified as an ethnic minority; 24.76% of students were first-generation, and 75.24% were continuing-

generation.

2.2. Procedure

The study adhered to the APA Ethical Standards and the Declaration of Helsinki guidelines and received ethical approval from the institutional review board of *Anonymized University*. The high school principals were contacted either by a research assistant or by undergraduate students enrolled in a thesis-preparatory credit course who also helped to carry out the recruitment process, the informed consent, and the data collection. High schools from Budapest and the countryside were selected using convenience sampling; however, we made a significant effort to gather data outside the capital and to obtain a socioeconomically diverse sample.

As the data collection took place in high schools, we needed to obtain consent from the principals and the class teachers first before contacting parents and students. As adolescents were aged between 14 and 21 years, the passive consent of parents was needed, meaning that they had to sign a form if they wanted to withdraw their children from the intervention. We also asked students for their informed consent. Once they had given their consent, students completed some basic demographics and were randomly assigned to either the treatment or the control condition. Randomization was within classes, as Qualtrics, the data gathering platform, randomly assigned each student to one of the two conditions.

The *intervention condition* was framed as contributing to an online media education program for parents and grandparents. The material introduced six scientifically supported strategies, all adapted from Guess et al. (2020), which were accompanied by peer testimonials to help participants identify online misinformation. The strategies included skepticism for headlines, looking beyond fear-mongering, inspecting the source of news, checking the evidence, triangulation, and considering if the story was a joke. These testimonials offered normative information on other students' negative experiences with failing to detect false information and also their favorable experiences when they managed to spot fake news (Walton & Cohen, 2011; Yeager et al., 2014; Yeager et al., 2016). We applied quotes for this purpose, abstaining from presenting descriptive norms in a numerical form (e.g., a percentage of students), unlike previous research (e.g., Yeager et al., 2014).⁷ After reading the strategies and the testimonials, students were asked to compose a concise letter addressed to a close, older family member, in which they would provide a summary of the six strategies and also share their thoughts on the most compelling arguments and guidance that could persuade their relatives to adopt these strategies into their daily routine. Participants in the treatment condition spent an average of 419 s (equivalent to 6.98 min) composing the letter and wrote a mean of 212.87 characters. The instructions, the psychological mechanisms of the self-persuasive message (Fig. S1), and a randomly chosen letter (Fig. S2) can be found online in Supplementary Materials.

Although the *control condition* closely resembled the intervention in its structure, it differed in that the control materials did not feature the topic of fake news. Instead, it was designed to guide the proper usage of social media platforms such as Facebook. The materials provided advice on how to avoid awkward behaviors exhibited by older generations that young adults may find uncomfortable. Additionally, the materials included six instances of implicit social norm violations, as reported by the participants (mixing up private messaging with Facebook's feed; sending virtual flowers on the wall for birthdays; incorrect use of emojis; uploading inappropriate profile pictures; anomalies during video chat; sending inappropriate invites to online games). Afterward, similar to the

⁶ The $d = 0.2$ is based on the US results of Guess et al. (2020).

⁷ We decided to abstain from presenting numerical descriptive norms as we did not have relevant statistics about Hungarian students at the time of conducting this intervention, and we did not want to provide fake data to participants.

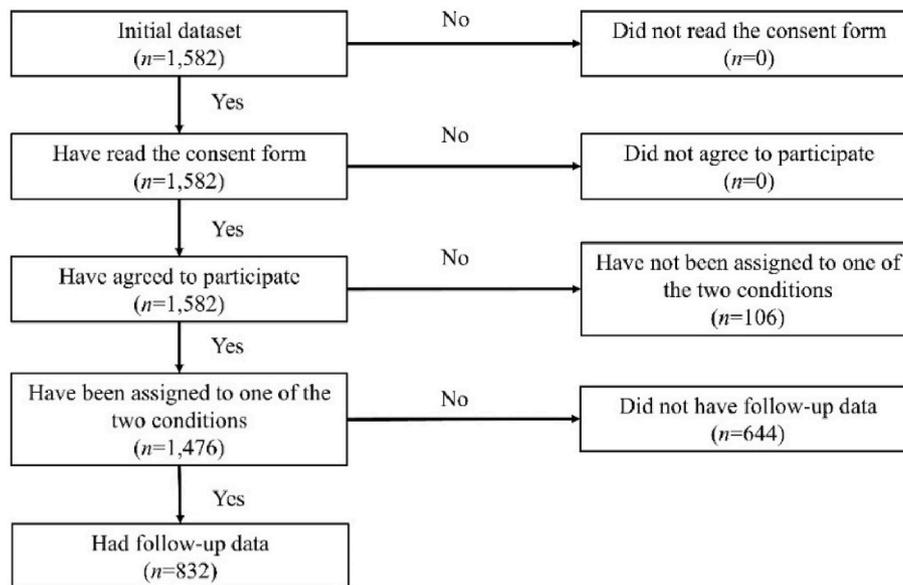


Fig. 1. Consort diagram.

treatment condition, participants were instructed to write a letter to an elderly family member outlining the key practices and offering advice on how to avoid negative behaviors on social media. Overall, the control condition focused on promoting positive social media behaviors without any mention of news content, fake news, or misinformation. The instructions of the control group can be found online in Supplementary Materials.

After writing the letters, students were presented with real and fake news items in a random order, and they had to evaluate the perceived accuracy of the headlines. At the end of the first data collection round, students could find out which news stories were true, and which were false. Four weeks later, they completed the follow-up test containing real and fake news items (different ones than in the first data collection round), and then they were debriefed for the real purpose of the study. The timeline of the study can be seen in Fig. 2.

2.3. Measures

2.3.1. Fake news evaluation accuracy. We applied Pennycook and Rand's (2019) method to evaluate the accuracy of fake and real news. Participants were presented with real and fake news items in a random order that had undergone thorough pretesting and were asked to rate their accuracy using four-point scales ("To the best of your knowledge, how accurate is the claim in the above headline?" *Not at all accurate/not very accurate/somewhat accurate/very accurate*).⁸ The fake and real news headlines differed in the post-intervention material and the follow-up (for the translated headlines, see the Supplemental Material). The reliability of these news items was acceptable (average inter-item correlation_{long-term} = 0.251, average inter-item correlation_{immediate} = 0.272, $ps < 0.001$). These values are acceptable considering that the headlines covered a wide range of topics and the scales did not measure one construct like personality scales; therefore, the lower reliability is not a

⁸ Following the criteria outlined by Pennycook, Epstein, et al. (2021), we chose real news articles from trustworthy Hungarian mainstream news sources. We abstained from choosing fact-checked real news because if real content needs to be fact-checked, this might carry ambiguity, potentially prompting readers to question its epistemological status. Conversely, we selected fake news from a webpage (Urbanlegends.hu), which contains fact-checked Hungarian fake news stories selected from different fake news sites (e.g., Titkokszigete.com).

cause for concern (see Faragó, Kende, & Krekó, 2020; Schmitt, 1996; Shnabel, Bar-Anan, Kende, Bareket, & Lazar, 2016, for further clarification). However, it is important to mention that some relevant studies found that confirming credible texts and questioning less credible texts derive from distinct cognitive processes (Fendt, Nistor, Scheibenzuber, & Artmann, 2023; Kiili, Räikkönen, Bråten, Strømsø, & Hagerman, 2023) and therefore alternative reliability measures are available.

2.3.1. Further individual difference variables for control and as possible moderators

Based on prior research on the relevance of cognitive abilities (D'Errico, Cicirelli, Corbelli, & Paciello, 2023; Faragó et al., 2023; Iacobucci & De Cicco, 2022; Leding & Antonio, 2019), digital literacy (Guess et al., 2020; Sirlin, Epstein, Arechar, & Rand, 2021), and socioeconomic variables (Duplaga, 2020; Georgiou, Delfabbro, & Balzan, 2020; Orosz et al., 2023) in recognizing misinformation, we preregistered the examination of the role of bullshit receptivity (five items, Pennycook, Cheyne, Barr, Koehler, & Fugelsang, 2015), cognitive reflection task (five items, Frederick, 2005 via Shenhav, Rand, & Greene, 2012; Thomson & Oppenheimer, 2016), need for cognition (five items, Cacioppo & Petty, 1982), digital literacy (five items, Guess & Munger, 2023), as control variables and as potential moderators. Socioeconomic status was measured by parental educational attainment (primary, secondary, and tertiary levels), specifically distinguishing between students without parents who hold a tertiary education degree and those who at least have one parent with a tertiary education degree. Ethnic minority status was assessed with one item ("Do you see yourself as belonging to an ethnic minority?"). A detailed description of these measures can be found in the online Supplemental Materials.

2.4. Analytic strategy

We employed multilevel linear regression (to account for the data's hierarchical structure, considering students as nested within schools) to assess the intervention's effect on fake news evaluation accuracy (H1a).⁹ Subsequently, we controlled for demographics (gender, ethnic minority

⁹ Originally, we aimed to apply OLS regressions following the pre-registration. However, after consulting with relevant experts, we modified this aspect of the preregistration, and applied multilevel linear regression analyses instead (students as nested within schools).

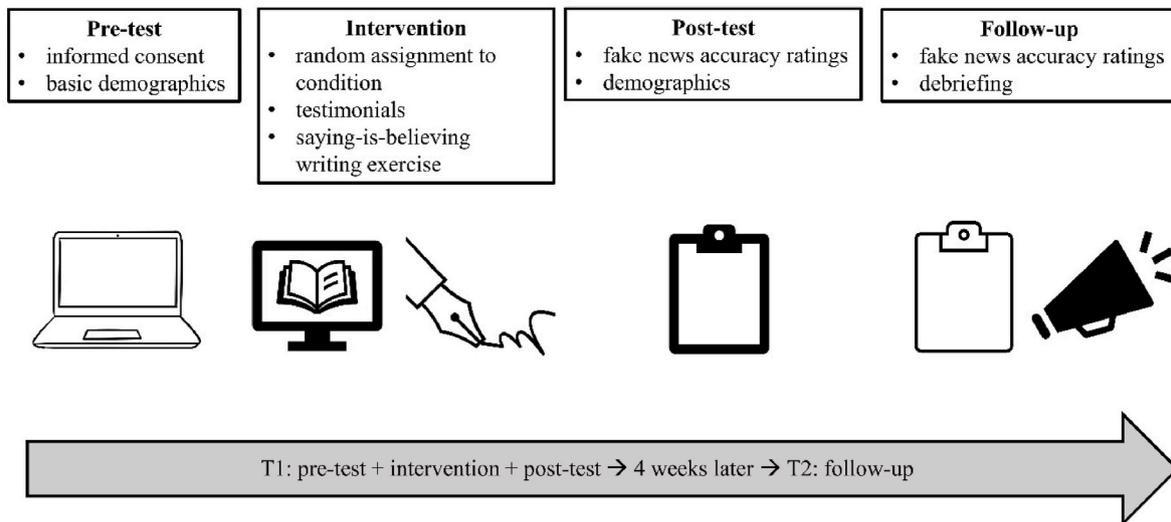


Fig. 2. Timeline.

status, and first-generation status) to examine the impact of the treatment independent of these individual differences. We performed these evaluations for the immediate (H1a) and the one-month follow-up (H1b).

We also employed multilevel linear regression analysis (students as nested within schools) to examine the interaction effect of moderators (analytic thinking, bullshit receptivity, digital literacy, need for cognition, socioeconomic status, and ethnic minority status) and intervention conditions on fake news evaluation accuracy ratings. Analytic thinking, bullshit receptivity, digital literacy, need for cognition, and subjective socioeconomic status were continuous variables, while first-gen status, ethnic minority status, and condition were treated as categorical moderators with two levels. Data was analyzed using R, version 4.3.0 (R Core Team, 2023).

Based on the preregistration, our primary analysis consists of an intent-to-treat analysis involving all participants who rated news items after the intervention. However, we cannot evaluate the data of students who withdrew before the outcome measures, as the appropriate dependent variables are unavailable. Additionally, we perform a treatment-of-the-treated analysis, examining the data of those students who actively followed the instructions in the intervention/control materials. We found this analysis useful as we were interested in whether the intervention can be effective among compliant adolescents.

For the treatment-of-the-treated analysis, we code the letters written by the students based on the following criteria. We included students if they (1) wrote about at least one strategy we provided in the instruction; (2) they did not simply copy the strategies; (3) they elaborated on a topic related to the task; (4) they wrote at least one sentence that included at least one strategy; (5) they gave a specific but not too broad advice (e.g., we dropped participants who wrote simply "be skeptical with the news" without more specific strategies); (6) they gave a piece of advice without the instructed letter format; (7) they advised regarding topics closely related to their condition: in the control condition, they gave at least one piece of advice regarding general digital literacy independently from the social networking context along the suggested dimensions; in the treatment group, they gave at least one piece of advice regarding fake news appearing on radio or in the television; (8) did not address the letter to a specific relative but more than one person or they addressed the pieces of advice to their relatives in general.

Following the development of coding schemes, we trained two independent raters to classify the letters using the pre-established coding categories. We made minor adjustments and clarifications to the coding system after an initial analysis of responses from 50 sample participants. This first phase of coding showed good consistency between raters, with

reliability scores of 0.74 ($CI: 0.706-0.772$) regarding the evaluation of students who followed vs. those who did not follow the instructions. Whenever there was a disagreement, the raters discussed the issue until they reached a consensus. In cases of unclear or overlapping statements, a third expert was consulted to resolve any ambiguities. All three coders collaboratively reviewed their coding methods and decisions to ensure consistent and accurate agreement.

After filtering out the compliant students, we conducted the same multilevel regression analyses in the treatment-of-the-treated analysis (as we do with the intent-to-treat sample), including the main effect and moderator effects.

3. Results

3.1. Preliminary analyses and attrition

We investigated the potential initial disparities between the conditions before the intervention. Subsequently, we examined the overall attrition (independent of the condition) and the attrition that varied depending on the condition in both samples (differential attrition). These analyses encompassed sociodemographic factors such as age, gender, ethnic minority status, parental level of education, type of settlement, and first-generation status, along with pertinent individual differences such as bullshit receptivity, need for cognition, and analytic thinking. Additionally, we explored the distinctions between students in the intent-to-treat sample and those in the treatment-of-the-treated sample.

To assess the random allocation of students to different conditions, we examined pre-intervention differences across the aforementioned psychological and sociodemographic variables. Our analysis revealed no statistically significant baseline differences between the two conditions (all $ps > 0.058$), providing evidence supporting the random condition assignment. Regarding the overall attrition, we observed a statistically significant decrease in the proportion of ethnic minority ($p < 0.001$), male ($p = 0.027$), and older students, as well as students with lower scores on cognitive reflection tasks ($p < 0.001$). However, we found no other notable differences in other sociodemographic and psychological variables between the participants who completed the follow-up and those who dropped out. Furthermore, concerning differential attrition, no statistically significant differences (all $ps > 0.088$) were observed between the students who were retained and those who were not retained across the two conditions.

We observed notable differences between adolescents who adhered to the instructions of the "saying-is-believing" exercise and those who

did not. Students who followed the instructions diligently were more likely to be female ($p < 0.001$), had parents with higher educational attainment ($p < 0.003$), exhibited a lower likelihood of being from minority backgrounds ($p < 0.001$), and possessed stronger analytic inclinations, as indicated by higher scores in analytic thinking ($p < 0.001$) and a greater need for cognition ($p < 0.001$), but they did not differ in terms of age ($p = 0.722$), type of settlement ($p = 0.868$), or bullshit receptivity ($p = 0.512$). However, due to the substantial similarity in the structure, demands, and even content regarding the intervention and control materials, students taking the instructions seriously were similar in the two conditions (all $ps > 0.078$). The only exception was gender: while there was a minimal difference between girls and boys taking the intervention condition seriously, girls took the control condition more seriously than boys ($ps < 0.001$). Lastly, we found no significant three-way interaction among condition, adherence to instructions, and participation in the long-term follow-up measures (all $ps > 0.231$). This implies that in both conditions, a comparable proportion of students took the instructions seriously and participated in the follow-up assessments. For an overview, see Fig. 3.

3.2. Intent-to-treat primary analyses

3.2.1. Short- and long-term fake news evaluation accuracy ratings ($H1_{ITT}$)

In general, the findings revealed that the intervention yielded statistically significant immediate ($b = 0.10$, $t(1453) = 3.30$, $k = 25$, $p < 0.001$, $d = 0.17$) but a non-significant long-term (one-month follow-up) ($b = 0.03$, $t(827) = 0.81$, $k = 25$, $p = 0.419$, $d = 0.05$) improvement in fake news evaluation accuracy ratings compared to the control condition. The immediate effects persisted even when accounting for pertinent individual differences in demographic variables, such as age, gender, parental education, and ethnic minority status ($b = 0.10$, $t(1439) = 3.25$, $k = 25$, $p < 0.001$, $d = 0.16$), while the long-term effects remained statistically insignificant when controlling for all the variables mentioned above ($b = 0.03$, $t(814) = 0.767$, $k = 25$, $p = 0.443$, $d = 0.05$).¹⁰

3.3. Intent-to-treat secondary analyses

Following the preregistration, an examination was conducted into the potential moderating effects of relevant socio-demographic variables (including age, gender, parental level of education, subjective socioeconomic status, and ethnic minority status) as well as individual difference variables (analytic thinking, bullshit receptivity, digital literacy, and need for cognition). Demographic variables ($p > 0.659$), digital literacy ($p = 0.489$), bullshit receptivity ($p = 0.411$), cognitive reflection ($p = 0.364$), and need for cognition ($p = 0.419$) did not yield statistically significant moderating effects on the treatment's impact on *immediate fake news evaluation accuracy scores*. Concerning the enduring effects of the intervention on *long-term fake news evaluation accuracy*, no demographic variables ($p > 0.054$), digital literacy ($p = 0.489$), bullshit receptivity ($p = 0.411$), cognitive reflection ($p = 0.137$), or need for cognition ($p = 0.103$) were identified as significant moderators.

3.4. Treatment-of-the-treated primary analyses

3.4.1. Short- and long-term fake news evaluation accuracy ratings ($H1_{TOT}$)

We carried out the same analyses by considering only those adolescents who took seriously the saying-is-believing, letter-writing task. Similarly to the intent-to-treat analyses, we found a significant main effect of the intervention on fake news evaluation accuracy ratings, ($b = 0.18$, $t(791) = 4.63$, $k = 25$, $p < 0.001$, $d = 0.29$) but a non-significant

long-term (one-month follow-up) ($b = 0.05$, $t(490) = 1.06$, $k = 25$, $p = 0.290$, $d = 0.09$) effect in fake news evaluation accuracy ratings compared to the control condition. Similarly to the intent-to-treat analyses, the immediate effects were present even when accounting for the above-mentioned demographic variables ($b = 0.17$, $t(787) = 4.32$, $k = 25$, $p < 0.001$, $d = 0.28$). Nevertheless, the long-term effects were not statistically significant when controlling for these demographic variables ($b = 0.04$, $t(488) = 0.767$, $k = 25$, $p = 0.361$, $d = 0.08$).

3.5. Treatment-of-the-treated secondary analyses

Following the preregistration, similarly to the intent-to-treat analyses, we examined the moderating effects of relevant socio-demographic variables (including age, gender, parental level of education, and ethnic minority status) as well as individual difference variables (analytic thinking, bullshit receptivity, digital literacy, and need for cognition). Demographic variables ($p > 0.488$), digital literacy ($p = 0.655$), bullshit receptivity ($p = 0.773$), cognitive reflection ($p = 0.293$), and need for cognition ($p = 0.509$) did not yield statistically significant moderating effects on the treatment's impact on *immediate fake news evaluation accuracy scores*. Concerning the enduring effects of the intervention on *long-term fake news evaluation accuracy*, no demographic variables ($p > 0.121$), digital literacy ($p = 0.540$), bullshit receptivity ($p = 0.415$), or cognitive reflection ($p = 0.233$) were significant moderators.

We found a significant moderation of the need for cognition on *long-term fake news evaluation accuracy* ($\beta = 0.23$, $t(488) = 2.54$, $k = 25$, $p = 0.011$). Adolescents with one standard deviation above the mean in need for cognition were more likely to recognize fake news in the intervention group compared to the control ($b = 0.16$, $t(488) = 2.54$, $k = 25$, $p = 0.012$, $d = 0.28$). These effects were even larger for those students who were two standard deviations above the mean in need for cognition ($b = 0.28$, $t(488) = 2.77$, $k = 25$, $p = 0.006$, $d = 0.51$). At the same time, the difference was not significant among students who were one ($p = .203$) or two ($p = .598$) standard deviations below the mean of need for cognition (see Fig. 4).

4. Discussion

In the current study, we aimed to implement a prosocial counter-misinformation wise intervention targeted at a relatively understudied population: adolescents. Within this intervention, participants were endowed with an expert role and were requested to write letters to digitally less proficient family members, outlining six strategies for identifying misinformation. The intervention resulted in a statistically significant immediate improvement ($d = 0.17$). Still, it did not show a significant long-term effect on fake news evaluation accuracy ratings compared to the control condition if we consider all adolescents. However, those adolescents who followed the instructions of the intervention and had a high need for cognition (one [$d = 0.28$] or two [$d = 0.51$] standard deviations above the mean) exhibited a significant advantage in accurately detecting fake news four weeks after the intervention. This pattern was not observed among students with a lower need for cognition.

The short-term effect size ($d = 0.17$) we obtained appears to be more modest than what Orosz et al. (2023) found applying their prosocial counter-misinformation intervention with a university sample ($d = 0.32$) or what Barzilai et al. (2023) reported for the short-term evaluation accuracy of inaccurate messages using gameplay ($d = 0.38$ for Study 2 but $d = 0.08$ for Study 1). However, these short-term effect sizes are similar to Guess et al.'s (2020) large-scale US digital literacy intervention study ($d = 0.20$) effects. Before the current intervention, no previous fake news interventions conducted among high school students measured long-term effects specifically on fake news evaluation accuracy. The long-term effects for engaged and highly analytical students ($d = 0.28$ and $d = 0.51$) are stronger than prior effects found with young

¹⁰ The intervention did not generate a general skepticism that would negatively affect the evaluation of real news (see details in the Supplemental Materials).

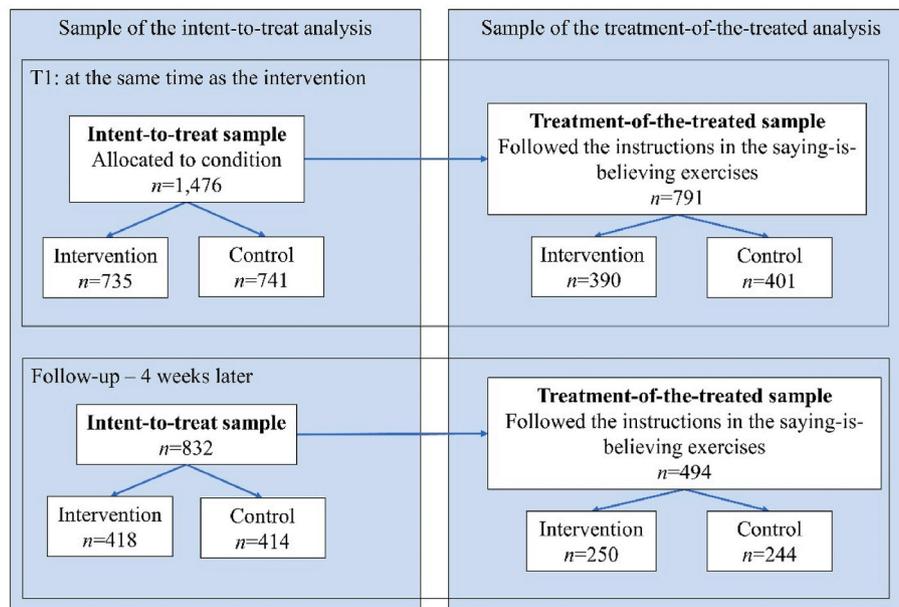


Fig. 3. Samples of the Intent-to-Treat and the Treatment-of-the-Treated analyses in the first data collection wave and the follow-up, split by conditions.

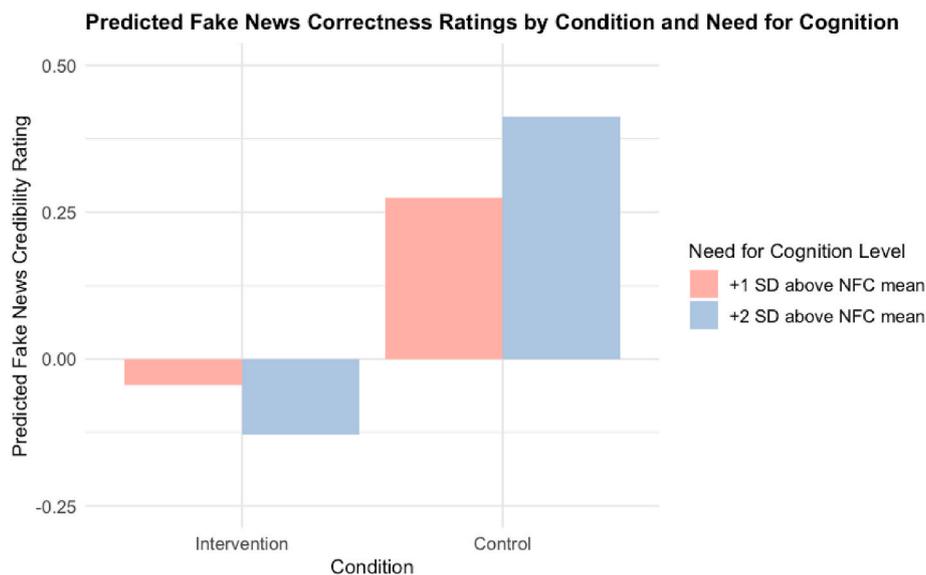


Fig. 4. Standardized and predicted fake news evaluation accuracy differences between the control and treatment groups among students who took the intervention task seriously and who are one and two standard deviations above the need for cognition mean.

Note. Low credibility values on the y-axis mean that adolescents find fake news less credible; higher values mean that they find fake news credible.

adult samples ($d = 0.22$, Orosz et al., 2023). Notably, the long-term effects we found for engaged and highly analytical individuals are much stronger than Guess et al. (2020) found as a long-term effect of their brief, online, scalable digital literacy intervention ($d = -0.08$). However, we must acknowledge that this large effect size was only observed among those who took the letter-writing task seriously and had a higher need for cognition, not the whole group of participants. Not only those were filtered out who did not pay attention to the letter-writing task, but we also had high attrition (43.6%) to the follow-up test due to the pandemic, which is a limitation of our study. Nevertheless, such long-term effect sizes have never been reported regarding a specific group of adolescents due to a 15-min-long online intervention.

An intervention can rarely have the same impact on everyone. In current social-psychological intervention studies (e.g., Walton et al.,

2023; Yeager et al., 2019), one of the most important questions is about who can and cannot benefit from interventions and identify those factors that can undermine the induction of recursive processes that can lead to beneficial long-term behavioral change (Walton & Yeager, 2020; Yeager & Walton, 2011). In our case, we identified two, not completely unexpected, factors that can undermine the success of the intervention. The first relates to the extent to which students take the intervention task seriously, while the second relates to their need for cognition. If students took it seriously and had a high need for cognition, the intervention led to solid long-term effects in misinformation recognition.

The moderating role of the need for cognition has been demonstrated in various interventions: for instance, highly analytic students benefitted most from substance use prevention programs (Giles et al., 2010) and from brief motivational interventions aimed at reducing alcohol consumption (Capone & Wood, 2009). Individuals with a higher need for

cognition also exhibited enhanced attitude-intention alignment following exposure to anti-doping messages (Horcajo, Santos, Guyer, & Moreno, 2019). They demonstrated increased self-efficacy for quitting smoking after participating in a written smoking cessation intervention (Haug et al., 2010). Regarding misinformation acceptance, some studies found that a higher need for cognition made individuals (with certain political orientations) less susceptible to fake information (e.g., Borah, 2022; Oliveira, 2023). The present study provides additional evidence supporting the positive impacts of engagement and motivation to think analytically in accurately recognizing misinformation.

The reasons adolescents may not take psychological interventions seriously can be multifaceted. Students' inclinations to follow the instructions of wise psychological interventions is a somewhat understudied topic (Schleider, Mullarkey, & Chacko, 2020). However, insights from the clinical and positive psychology literature (Aguirre Velasco, Cruz, Billings, Jimenez, & Rowe, 2020; Bolier et al., 2013; Radez et al., 2021) suggest that barriers to seeking professional help for anxiety and depression can be related to shame, stigma, and a lack of knowledge about mental health problems. In addition, their desire for autonomy and independence can also prevent adolescents from participating in these clinical programs. Additionally, a systematic review of quantitative and qualitative studies found that a lack of mental health knowledge and negative perceptions of help-seeking were common reasons for not seeking professional help. Furthermore, according to the clinical literature, the balance between a teenager's autonomy and adult input is delicate, and interventions need to be developed with the input of young people to meet their needs (see Aguirre Velasco et al., 2020; Radez et al., 2021). In the present case, special attention was paid to designing autonomy-supportive intervention material, allowing adolescents to find themselves in a superior, expert position to give advice. However, future research is needed to explore how to get *all* students better invested in the intervention.

There is one more aspect that can be improved in future studies. The authors were not present personally when the interventions were deployed in the classroom, and the teachers requested students to participate. Even though we provided them with a framing, it is possible that they did not always follow the framing we provided. Their framing might have generated resistance. In future studies, to increase engagement with the intervention materials, it is possible to assign this instructional role to someone who is not the teacher but an independent individual who is blind to the real purpose of the intervention and the conditions.

4.1. Applied and theoretical contributions

Our intervention study addresses four main challenges in the fake news intervention literature. The first one is related to the *nature of the intervention itself*: previous attempts tried to make people more vigilant using accuracy-nudging interventions (Fazio, 2020; Pennycook et al., 2020, 2021; Pennycook & Rand, 2021), taught them digital literacy skills, or inoculated them with the common tricks that fake news creators apply (Basol et al., 2021; Guess et al., 2020; Kahne & Bowyer, 2017; Maertens et al., 2020, 2021; Maertens et al., 2023, preprint; Roozenbeek & Van der Linden, 2019; Roozenbeek, Van der Linden, Goldberg, Rathje, & Lewandowsky, 2022; Roozenbeek et al., 2020; Scheibenzuber, Hofer, & Nistor, 2021; Van der Linden, Roozenbeek, & Compton, 2020; Zerback et al., 2020), so all of them tried to teach people *how* to spot misinformation. However, our prosocial counter-misinformation intervention also gives reasons *why* fake news deception is important. Therefore, unlike prior attempts, our intervention facilitates meaning-making processes (Walton & Wilson, 2018). Research shows that prosocial motivations increase helping behavior in different social contexts (see e.g., Grant, 2007; Grant & Hofmann, 2011; Grant & Shandell, 2022; Paunesku et al., 2015; Reeves et al., 2021; Yeager et al., 2014), and our results suggest that these motives (aligned with the expert role and digital mindset) can also help decrease the perceived

accuracy of fake information.

The second challenge relates to the *scalability* of misinformation interventions. Though this prosocial intervention lasts approximately 15 min (including the testimonials and the letter-writing task), it is less concise than accuracy-nudging interventions (Pennycook et al., 2020, 2021; Pennycook & Rand, 2021). However, it is less time-consuming than many competency-building or inoculation interventions. For instance, participating in competency-fostering courses (Caroti et al., 2023, preprint; Scheibenzuber et al., 2021) or playing games (Basol et al., 2021; Basol, Roozenbeek, & van der Linden, 2020; Roozenbeek et al., 2020; Roozenbeek & Van der Linden, 2019; Van der Linden et al., 2020) lasts for from half an hour to several hours (Pennycook & Rand, 2021), which further complicates the large-scale adoption of these interventions despite their effectiveness. Nevertheless, digital media literacy interventions using tips for identifying fake news on social media (Guess et al., 2020), short games (Basol et al., 2021), or brief video inoculations (Roozenbeek et al., 2022) are promising and scalable approaches to counter misinformation. Our approach demonstrates the same potential for scalability – at least in a classroom context –, as it harnesses efficient delivery methods while retaining the capacity for meaningful and lasting impact. However, future research is needed to investigate the applicability of the present intervention in various online contexts, that is, whether it can be distributed on social media as efficiently as accuracy-nudging interventions (Pennycook, Epstein, et al., 2021) or short video inoculations (Roozenbeek et al., 2022).

The third challenge pertains to the *enduring impact* of interventions targeting fake news. Despite the pressing need for sustained approaches that effectively reduce the perceived credibility of false information, research examining the long-term effects of such interventions remains limited (or even lacking, e.g., for accuracy-nudging interventions). Previous studies have predominantly focused on digital media literacy and inoculation interventions, yet the observed effects have been relatively short-lived, ranging from one week (Basol et al., 2021; Maertens et al., 2020) to three weeks (Guess et al., 2020). Only a few interventions have demonstrated enduring outcomes beyond 30 days without implementing additional reinforcement measures (Caroti et al., 2023, preprint; Maertens et al., 2023, preprint, Study 1 and Study 5; Orosz et al., 2023). In this context, our intervention not only exhibits scalability but also engenders a sustained reduction in fake news evaluation accuracy among engaged and highly analytical students, as evidenced by the statistically significant effects persisting for four weeks post-intervention without the need for booster sessions.

The fourth issue pertains to the limited exploration of the *adolescent age group* in misinformation intervention research. Despite their heightened vulnerability to fake news (Baptista et al., 2022; Common Sense Media, 2017; Papapicco et al., 2022), only a limited number of intervention studies have targeted adolescents so far (Barzilai et al., 2023; Caroti et al., 2023, in press; Delaney et al., 2022; Smith & de los Santos, 2022). The present randomized controlled study contributes to the literature by directly addressing susceptibility to fake news among high school students and examining the intervention's short- and long-term efficacy, distinguishing our approach from previous research endeavors. Additionally, we implemented a more extensive data collection approach, gathering a significantly larger sample size than previous interventions targeted at high school students (refer to Barzilai et al., 2023; Caroti et al., 2023, in press). Furthermore, our study is preregistered, and data was collected from a non-WEIRD¹¹ country. Being a 15-min-long brief online intervention, no external support is required for reading the materials or composing the letters, making it easily integrated into the high school curriculum.

In sum, the present work provides a possible answer to these challenges. First, the present intervention provided a reason for why it might

¹¹ The abbreviation stands for Western, Educated, Industrialized, Rich, and Democratic.

be relevant to spot misinformation and not only gave some advice about how they can do it. Second, reading the testimonials and writing the letter takes 15 min on average; therefore, it is potentially scalable in the classroom context. Third, we found small-to-moderate long-term effects regarding certain students (and not among all). Fourth, unlike most prior intervention attempts, the present focuses on adolescents. Besides demonstrating how the present work can answer these four current challenges of the field, it is similarly important to list its limitations.

4.2. Limitations and future directions

Our study has several limitations that warrant consideration. First, we could not compensate participating students due to legal constraints. This lack of compensation may have influenced the quality of their responses, the attrition rate, and their overall motivation.

Second, our data collection occurred during the fourth wave of the COVID-19 pandemic, which likely impacted the attrition rate between the initial and subsequent data collection rounds. The pandemic's unique circumstances may have introduced confounding variables into our study.

Third, unfortunately, Hungary's National Public Education Act imposes restrictions on school training concerning politically sensitive topics such as prejudice reduction and LGBTQ + rights support ([National Public Education Act, 2011, 9/A. §](#)). Consequently, despite being soon-to-be voters, we could not administer political fake news stimuli to this age group. This limitation hampers the comprehensive examination of the effects of political misinformation on our sample.

Fourth, although we did our best to design similarly demanding and engaging active control materials, we found that female students took the control group exercise more seriously than male students. However, in a country where gender stereotypes are strong ([Kántás, Faragó, & Kovacs, 2022](#)), it is not easy to identify an online control group topic that is equally interesting and engaging to boys and girls. Next time, we should change the control material to become more appealing to boys, making them take it more seriously.

Fifth, it is important to note that the significant long-term effect was observed solely among participants who approached the letter-writing task with seriousness and exhibited a higher need for cognition, not the whole sample of students. Not only were adolescents who did not engage diligently with the task excluded, but we also encountered a significant attrition rate (43.6%) during the follow-up test, which represents a limitation of our study. Future research should investigate strategies to foster greater engagement among all students in the intervention.

The sixth limitation concerns the mechanisms that drive the intervention's effects. We did not explore mediators (prosociality, digital mindset, expert role), as our primary objective was to identify potential long-term effects and relevant moderators influencing the effectiveness of the intervention. Future studies can identify the role of these psychological mechanisms through manipulation checks and the specific role of three mechanisms in the immediate and long-term effectiveness of the intervention.

The seventh limitation is related to the underlying cognitive processes. Although Pennycook and Rand (2020) found a unidimensional structure of the fake and real news suggesting that a single cognitive structure might be responsible for processing fake and real news, [Kiili et al. \(2023\)](#) and [Fendt et al. \(2023\)](#) identified that two underlying cognitive structures are responsible for evaluating real and fake news. Based on supplemental analyses (see Supplemental Materials) the present data does not belong unambiguously to either the unifactorial or the two-factor solutions. Therefore, future studies might investigate this aspect in the Hungarian context.

The eighth limitation is that the mean and the standard deviation of the accuracy ratings of the fake news items were lower in the follow-up than immediately after the intervention. This might indicate that students could recognize them more easily independently from their

condition, which could also contribute to the less salient long-term effects of the intervention. In future studies, it might be useful not only to randomly present fake and real news within each measurement point (within the post-test and the follow-up) but to mix these items between the measurement points (between the post-test and the follow-up).

The ninth limitation is related to the low reliability of the outcome variable: despite a wide range of topics in the news materials with relatively low reliability does not a cause for concern based on prior studies (see [Faragó et al., 2020](#); [Schmitt, 1996](#); [Shnabel et al., 2016](#), for further clarification). However, if the goal is respecting high psychometric standards and an item set with high-reliability scores, it might be useful to implement items from the same topic (e.g., medical news related to sugar consumption or COVID-19; see [Kiili et al. \(2023\)](#) and [Fendt et al. \(2023\)](#)).

The tenth limitation is about measuring more outcome variables: even though our intervention made adolescents more aware of the presence of fake news, this does not automatically imply they possess sufficient information to decide whether a news article is real or fake, as they may still need to pursue further information to assess the extent of falsehood within a piece of news. Therefore, future studies should include other dependent variables than fake news evaluation accuracy, like the intention to seek additional information.

Finally, although the news items were pretested on a comprehensive adult sample, we did not have a similarly extended pretest among adolescents. Furthermore, we did not have a pre-intervention fake and real news accuracy assessment measure to provide a pre-intervention baseline in fake accuracy ratings. The main reason for not doing so was the time limit (the intervention material and the pre-and post-measures should fit in a standard classroom lesson), and we did not aim to prime students with an intellectual task before the intervention that can serve as a potential confounding accuracy nudge.

5. Conclusion

Our study provided evidence for the short- and some long-term effects of a scalable prosocial fake news intervention among high school students. Though the intervention was effective for the whole treatment group in the short term, motivated participants who elaborated on the task, took the letter-writing seriously, and had a higher need for cognition had a more robust long-term decrease in fake news evaluation accuracy. Our intervention method incorporates indirect persuasive techniques that have been shown to facilitate enduring changes in attitudes and behaviors ([Aronson, 1999](#)) in ideal circumstances. However, similar to other wise social psychological interventions ([Walton et al., 2023](#); [Yeager et al., 2019](#)), adolescents' motivation to engage in the intervention is indeed a crucial prerequisite for the desired effects. Though adolescents were endowed with an expert role in the digital world rather than treated as incompetent teenagers, and they were given respectful treatment, which is a requirement to increase adolescents' engagement and motivation in interventions ([Yeager et al., 2018](#)), this might not be enough for all adolescents to be motivated. Besides overcoming the four challenges we sketched in the discussion, future studies might explore factors that could enhance motivation and the willingness to elaborate a counter-misinformation intervention content among this age group.

Prior dissemination of the data

As our study was funded by the European Union's Horizon 2020 research and innovation program (grant agreement No 822590), we had to disseminate the preliminary findings in the form of a report, which can be found here: https://openarchive.tk.mta.hu/555/1/tas:k8.4-b-fn_iv_02_15-M_kp-28-04-final.pdf.

CRedit authorship contribution statement

Gábor Orosz: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Laura Faragó:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Benedek Paskuj:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Péter Krekó:** Writing – review & editing, Writing – original draft, Funding acquisition.

Declaration of generative AI and AI-assisted technologies in the writing process

While preparing this work, we used ChatGPT and Grammarly to check grammar and spelling to increase the quality and fluency of the manuscript. After using them, we reviewed and edited the content as needed and take full responsibility for the publication's content.

The preregistered hypotheses and analysis plan can be found here: <https://osf.io/8tqk6>.

Declaration of competing interest

We have no conflict of interest to declare.

Data availability

Our data file, the analysis codes, and the materials are also openly available on Open Science Framework (DOI: 10.17605/OSF.IO/EG2VS) and can be found here: <https://osf.io/eg2vs/>.

Acknowledgments

This research was financed in the framework of the DEMOS project. This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 822590. The consortium leader was Zsolt Boda. For more details, see <https://demos-h2020.eu/en>. Laura Faragó was supported by ÚNKP-23-4 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund, grant nr.: ÚNKP-23-4. Gábor Orosz was supported by Northern French Strategic Dialogue (Phase 1 & 2), STARS, and CPJ grants. Special thanks to the Schmidt Foundation postdoc grant that made it possible to study prosocial wise interventions at Stanford University. Special thanks to Williams Nuytens and Pasquale Mammone for their institutional support. Péter Krekó was supported by the János Bolyai Research Fellowship of the Hungarian Academy of Sciences (grant nr: BO/00686/20/2) and by the Hungarian National Research, Development and Innovation Office (grant nr: K147292). We would also like to thank all the BA students for their valuable help in facilitating recruitment, obtaining informed consent, collecting data in high schools, and coding letters.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.chb.2024.108338>.

References

Aguirre Velasco, A., Cruz, I. S. S., Billings, J., Jimenez, M., & Rowe, S. (2020). What are the barriers, facilitators and interventions targeting help-seeking behaviours for common mental health problems in adolescents? A systematic review. *BMC Psychiatry*, 20(1), 1–22. <https://doi.org/10.1186/s12888-020-02659-0>

- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *The Journal of Economic Perspectives*, 31, 211–236. <https://doi.org/10.1257/jep.31.2.211>
- Allington, D., Duffy, B., Wessely, S., Dhavan, N., & Rubin, J. (2021). Health-protective behavior, social media usage and conspiracy belief during the COVID-19 public health emergency. *Psychological Medicine*, 51, 1763–1769. <https://doi.org/10.1017/S003329172000224X>
- Aronson, E. (1999). The power of self-persuasion. *American Psychologist*, 54, 875–884. <https://doi.org/10.1037/h0088188>
- Aronson, E., Fried, C., & Stone, J. (1991). Overcoming denial and increasing the intention to use condoms through the induction of hypocrite. *American Journal of Public Health*, 81, 1636–1638. <https://doi.org/10.2105/AJPH.81.12.1636>
- Balmas, M. (2014). When fake news becomes real: Combined exposure to multiple news sources and political attitudes of inefficacy, alienation, and cynicism. *Communication Research*, 41, 430–454. <https://doi.org/10.1177/009365021245360>
- Baptista, J. P., Gradim, A., & Correia, E. (2022). A relação entre a crença em fake news e as estratégias para procurar informação dos jovens portugueses. *Observatorio*, 16, 203–232.
- Barzilai, S., Mor-Hagani, S., Abed, F., Tal-Savir, D., Goldik, N. A., Talmon, I., et al. (2023). Misinformation Is Contagious: Middle school students learn how to evaluate and share information responsibly through a digital game. *Computers & Education*, 202, Article 104832. <https://doi.org/10.1016/j.compedu.2023.104832>
- Basol, M., Roozenbeek, J., Berriche, M., Uenal, F., McClanahan, W. P., & van der Linden, S. (2021). Towards psychological herd immunity: Cross-cultural evidence for two prebunking interventions against COVID-19 misinformation. *Big Data & Society*, 8(1), Article 20539517211013868. <https://doi.org/10.1177/20539517211013868>
- Basol, M., Roozenbeek, J., & van der Linden, S. (2020). Good news about bad news: Gamified inoculation boosts confidence and cognitive immunity against fake news. *Journal of Cognition*, 3, 2. <https://doi.org/10.5334/joc.91>
- Benaissa Pedriza, S. (2021). Sources, channels, and strategies of disinformation in the 2020 US election: Social networks, traditional media, and political candidates. *Journalism and Media*, 2(4), 605–624. <https://doi.org/10.3390/journalmedia2040036>
- Bolier, L., Haverman, M., Westerhof, G. J., Riper, H., Smit, F., & Bohlmeijer, E. (2013). Positive psychology interventions: A meta-analysis of randomized controlled studies. *BMC Public Health*, 13(1), 1–20. <https://doi.org/10.1186/1471-2458-13-119>
- Borah, P. (2022). The moderating role of political ideology: Need for cognition, media locus of control, misinformation efficacy, and misperceptions about COVID-19. *International Journal of Communication*, 16, 3534–3559.
- Bozóki, A., & Hegedűs, D. (2018). An externally constrained hybrid regime: Hungary in the European union. *Democratization*, 25, 1173–1189. <https://doi.org/10.1080/13510347.2018.1455664>
- Brashier, N. M., & Schacter, D. L. (2020). Aging in an era of fake news. *Current Directions in Psychological Science*, 29(3), 316–323. <https://doi.org/10.1002/asi.23784>
- Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42(1), 116–131. <https://doi.org/10.1037/0022-3514.42.1.116>
- Capone, C., & Wood, M. D. (2009). Thinking about drinking: Need for cognition and readiness to change moderate the effects of brief alcohol interventions. *Psychology of Addictive Behaviors*, 23(4), 684–688. <https://doi.org/10.1037/a0016235>
- Caroti, D., Adam-Troian, J., Bagneux, V., & Theraud, M. (2023). Critical thinking education to decrease conspiracy and paranormal beliefs among secondary school students: A phase I trial. Preprint <https://doi.org/10.31234/osf.io/p5qzg>.
- Common Sense Media. (2017). New report: Kids value the news media but feel misrepresented, neglected by coverage. Retrieved from <https://www.common sense media.org/press-releases/new-report-kids-value-the-news-media-but-feel-misrepresented-neglected-by-coverage>. (Accessed 6 January 2024).
- Corbu, N., Oprea, D. A., & Frunzaru, V. (2022). Romanian adolescents, fake news, and the third-person effect: A cross-sectional study. *Journal of Children and Media*, 16, 387–405. <https://doi.org/10.1080/17482798.2021.1992460>
- Csöte, A. (2009). Boldogtalan kapitalizmus? A mai magyarországi társadalom értékpreferenciáinak néhány jellemzője. In Á. Szalai (Ed.), *Kapitalista elvárások. Közjő és kapitalizmus intézet*.
- 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*, 1–20. <https://doi.org/10.1007/s11218-023-09777-z>
- Delaney, C. J., Bacon, H. R., & Matson, S. F. (2022). From fake news to racism: A study of change in a reading intervention class. *Journal of Adolescent & Adult Literacy*, 65, 419–429. <https://doi.org/10.1002/jaal.1223>
- Demeter, M. (2018). Propaganda against the West in the Heart of Europe. A masked official state campaign in Hungary. *Central European Journal of Chemistry*, 11, 177–197. [https://doi.org/10.19195/1899-5101.11.2\(21\).5](https://doi.org/10.19195/1899-5101.11.2(21).5)
- Duplaga, M. (2020). The determinants of conspiracy beliefs related to the COVID-19 pandemic in a nationally representative sample of internet users. *International Journal of Environmental Research and Public Health*, 17(21), 7818. <https://doi.org/10.3390/ijerph17217818>
- Faragó, L., Kende, A., & Krekó, P. (2020). We only believe in news that we doctored ourselves: The connection between partisanship and political fake news. *Social Psychology*, 51, 77–90. <https://doi.org/10.1027/1864-9335/a000391>
- Faragó, L., Krekó, P., & Orosz, G. (2023). Hungarian, lazy, and biased: The role of analytic thinking and partisanship in fake news discernment on a Hungarian representative sample. *Scientific Reports*, 13, 178. <https://doi.org/10.1038/s41598-022-26724-8>
- Fazio, L. (2020). Pausing to consider why a headline is true or false can help reduce the sharing of false news. *Harvard Kennedy School Misinformation Review*. <https://doi.org/10.37016/mr-2020-009>

- Fendt, M., Nistor, N., Scheibenzuber, C., & Artmann, B. (2023). Sourcing against misinformation: Effects of a scalable lateral reading training based on cognitive apprenticeship. *Computers in Human Behavior*, 146, Article 107820. <https://doi.org/10.1016/j.chb.2023.107820>
- Frederick, S. (2005). Cognitive reflection and decision making. *The Journal of Economic Perspectives*, 19(4), 25–42. <https://doi.org/10.1257/089533005775196732>
- Freedom House. (2019). *Freedom in the world 2019*. Hungary. Retrieved from <https://freedomhouse.org/country/hungary/freedom-world/2019>. (Accessed 6 January 2024).
- Georgiou, N., Delfabbro, P., & Balzan, R. (2020). COVID-19-related conspiracy beliefs and their relationship with perceived stress and pre-existing conspiracy beliefs. *Personality and Individual Differences*, 166, Article 110201. <https://doi.org/10.1016/j.paid.2020.110201>
- Giles, S. M., Pankratz, M. M., Ringwalt, C., Hansen, W. B., Dusenbury, L., & Jackson-Newsom, J. (2010). Teachers' delivery skills and substance use prevention program outcomes: The moderating role of students' need for cognition and impulse decision making. *Journal of Drug Education*, 40(4), 395–410. <https://doi.org/10.2190/DE.40.4.e>
- Grant, A. M. (2007). Relational job design and the motivation to make a prosocial difference. *Academy of Management Review*, 32(2), 393–417. <https://doi.org/10.5465/amr.2007.24351328>
- Grant, A. M., & Hofmann, D. A. (2011). It's not all about me: Motivating hand hygiene among health care professionals by focusing on patients. *Psychological Science*, 22(12), 1494–1499. <https://doi.org/10.1177/0956797611419172>
- Grant, A. M., & Shandell, M. S. (2022). Social motivation at work: The organizational psychology of effort for, against, and with others. *Annual Review of Psychology*, 73, 301–326. <https://doi.org/10.1146/annurev-psych-060321-033406>
- Guess, A. M., Lerner, M., Lyons, B., Montgomery, J. M., Nyhan, B., Reifler, J., et al. (2020). A digital media literacy intervention increases discernment between mainstream and false news in the United States and India (Vol. 117, pp. 15536–15545). Proceedings of the National Academy of Sciences. <https://doi.org/10.1073/pnas.1920498117>
- Guess, A. M., & Munger, K. (2023). Digital literacy and online political behavior. *Political Science Research and Methods*, 11(1), 110–128. <https://doi.org/10.1017/psrm.2022.17>
- Guriev, S., & Treisman, D. (2022). *Spin dictators: The changing face of tyranny in the 21st century*. Princeton University Press.
- Hankiss, E. (1990). *East European alternatives*. USA: Oxford University Press.
- Haug, S., Meyer, C., Ulbricht, S., Gross, B., Rumpf, H. J., & John, U. (2010). Need for cognition as a predictor and a moderator of outcome in a tailored letters smoking cessation intervention. *Health Psychology*, 29(4), 367–373. <https://doi.org/10.1037/a0019450>
- 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, 1–27. <https://doi.org/10.1080/11356405.2020.1859739>
- Higgins, E. T., & Rholes, W. S. (1978). "Saying is believing": Effects of message modification on memory and liking for the person described. *Journal of Experimental Social Psychology*, 14(4), 363–378. [https://doi.org/10.1016/0022-1031\(78\)90032-X](https://doi.org/10.1016/0022-1031(78)90032-X)
- Horcajo, J., Santos, D., Guyer, J. J., & Moreno, L. (2019). Changing attitudes and intentions related to doping: An analysis of individual differences in need for cognition. *Journal of Sports Sciences*, 37(24), 2835–2843. <https://doi.org/10.1080/02640414.2019.1665876>
- Iacobucci, S., & De Cicco, R. (2022). A literature review of bullshit receptivity: Perspectives for an informed policy-making against misinformation. *Journal of Behavioral Economics for Policy*, 6, 23–40.
- Kántás, É. M., Faragó, L., & Kovacs, M. (2022). If you can dream it, you can do it! – The role of sexual orientation in preferences toward boys' and girls' career orientation and gendered behavior. *European Journal of Social Psychology*, 52(2), 305–325. <https://doi.org/10.1002/ejsp.2789>
- Kahne, J., & Bowyer, B. (2017). Educating for democracy in a partisan age: Confronting the challenges of motivated reasoning and misinformation. *American Educational Research Journal*, 54, 3–34. <https://doi.org/10.3102/000283121679817>
- Kiili, C., Rääkkönen, E., Bråten, I., Strømsø, H. I., & Hagerman, M. S. (2023). Examining the structure of credibility evaluation when sixth graders read online texts. *Journal of Computer Assisted Learning*, 39(3), 954–969. <https://doi.org/10.1111/jcal.12779>
- Krekó, P. (2022). The birth of an illiberal information autocracy in Europe: A case study on Hungary. *The Journal of Illiberalism Studies*, 2, 55–72. <https://doi.org/10.53483/WCJW3538>
- Krekó, P., & Enyedi, Z. (2018). Explaining Eastern Europe: Orbán's laboratory of illiberalism. *Journal of Democracy*, 29, 39–51. <https://doi.org/10.1353/jod.2018.0043>
- LaCasse, J., Canning, E. A., Bowman, N. A., Murphy, M. C., & Logel, C. (2020). A social-belonging intervention improves STEM outcomes for students who speak English as a second language. *Science Advances*, 6(40), 1–10. <https://doi.org/10.1126/sciadv.abb6543>
- Leding, J. K., & Antonio, L. (2019). Need for cognition and discrepancy detection in the misinformation effect. *Journal of Cognitive Psychology*, 31(4), 409–415. <https://doi.org/10.1080/20445911.2019.1626400>
- Lewandowsky, S., Stritzke, W. G., Freund, A. M., Oberauer, K., & Krueger, J. I. (2013). Misinformation, disinformation, and violent conflict: From Iraq and the "War on Terror" to future threats to peace. *American Psychologist*, 68, 487–501. <https://doi.org/10.1037/a0034515>
- Maertens, R., Anseel, F., & van der Linden, S. (2020). Combatting climate change misinformation: Evidence for longevity of inoculation and consensus messaging effects. *Journal of Environmental Psychology*, 70, Article 101455. <https://doi.org/10.1016/j.jenvp.2020.101455>
- Maertens, R., Roozenbeek, J., Basol, M., & van der Linden, S. (2021). Long-term effectiveness of inoculation against misinformation: Three longitudinal experiments. *Journal of Experimental Psychology: Applied*, 27, 1–16. <https://doi.org/10.1037/xap0000315>
- Maertens, R., Roozenbeek, J., Simons, J., Lewandowsky, S., Maturo, V., Goldberg, B., et al. (2023). Psychological booster shots targeting memory increase long-term resistance against misinformation. <https://doi.org/10.31234/osf.io/6r9as>.
- Marchi, R. (2012). With Facebook, blogs, and fake news, teens reject journalistic "objectivity". *Journal of Communication Inquiry*, 36, 246–262. <https://doi.org/10.1177/0196859912458700>
- National Public Education Act. (2011). (Cth).
- Newman, N., Fletcher, R., Kalogeropoulos, A., & Nielsen, R. K. (2019). *Reuters institute digital news report 2019*. <http://www.digitalnewsreport.org/survey/2019/overview-key-findings-2019/>.
- Oliveira, C. (2023). Misinformation, political preferences, and cognitive traits: A look at the Brazilian electorate. *Opinião Pública*, 29, 304–326. <https://doi.org/10.1590/1807-01912023292304>
- Orosz, G., Paskuj, B., Faragó, L., & Krekó, P. (2023). A prosocial fake news intervention with durable effects. *Scientific Reports*, 13, 3958. <https://doi.org/10.1038/s41598-023-30867-7>
- Pátkainé Bende, A. (2022). *Értékek, attitűdök és vélemények a családról a European Values Study 2017-es eredményeinek tükrében* Kapocs, 3–4, 13–22.
- 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, 20. <https://doi.org/10.3390/mti6030020>
- Paul, J., Macedo-Rouet, M., Rouet, J.-F., & Stadler, M. (2017). Why attend to source information when reading online? The perspective of ninth grade students from two different countries. *Computers & Education*, 113, 339–354. <https://doi.org/10.1016/j.compedu.2017.05.020>
- Paunescu, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., & Dweck, C. S. (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological Science*, 26(6), 784–793. <https://doi.org/10.1177/0956797615571017>
- Pennycook, G., Binnendyk, J., Newton, C., & Rand, D. G. (2021). A practical guide to doing behavioral research on fake news and misinformation. *Collabra: Psychology*, 7(1), Article 25293. <https://doi.org/10.1525/collabra.25293>
- Pennycook, G., Cheyne, J. A., Barr, N., Koehler, D. J., & Fugelsang, J. A. (2015). On the reception and detection of pseudo-profound bullshit. *Judgment and Decision Making*, 10(6), 549–563. <https://doi.org/10.1017/S1930297500006999>
- Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature*, 592, 590–595. <https://doi.org/10.1038/s41586-021-03344-2>
- Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science*, 31, 770–780. <https://doi.org/10.1177/0956797620939054>
- 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. <https://doi.org/10.1016/j.cognition.2018.06.011>
- Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, 25, 388–402. <https://doi.org/10.1016/j.tics.2021.02.007>
- Political Capital. (2023). Disinformation wonderland in the Hungarian government-controlled online media: Origo's articles on Putin and Zelensky. Retrieved from https://politicalcapital.hu/news.php?article_read=1&article_id=3192. (Accessed 6 January 2024).
- Polyák, G., Urbán, Á., & Szávai, P. (2022). Information patterns and news bubbles in Hungary. *Media and Communication*, 10(3), 133–145. <https://doi.org/10.17645/mac.v10i3.5373>
- R Core Team. (2023). *R: A language and environment for statistical computing* Version 4.3.0. <https://www.R-project.org/>.
- Radez, J., Reardon, T., Creswell, C., Lawrence, P. J., Evdoka-Burton, G., & Waite, P. (2021). *Why do children and adolescents (not) seek and access professional help for their mental health problems? A systematic review of quantitative and qualitative studies* (Vol. 30, pp. 183–211). European Child & Adolescent Psychiatry. <https://doi.org/10.1007/s00787-019-01469-4>
- Reeves, S. L., Henderson, M. D., Cohen, G. L., Steingut, R. R., Hirschi, Q., & Yeager, D. S. (2021). Psychological affordances help explain where a self-transcendent purpose intervention improves performance. *Journal of Personality and Social Psychology*, 120(1), 1–15. <https://doi.org/10.1037/pspa0000246>
- Reporters Without Borders. (2009). Press freedom index: Global score. Retrieved from <https://rsf.org/en/index?year=2009>. (Accessed 6 January 2024).
- Reporters Without Borders. (2022). Press freedom index: Hungary. Retrieved from <https://rsf.org/en/country/hungary>. (Accessed 6 January 2024).
- Roozenbeek, J., & Van der Linden, S. (2019). Fake news game confers psychological resistance against online misinformation. *Palgrave Communications*, 5, 1–10. <https://doi.org/10.1057/s41599-019-0279-9>
- Roozenbeek, J., Van der Linden, S., Goldberg, B., Rathje, S., & Lewandowsky, S. (2022). Psychological inoculation improves resilience against misinformation on social media. *Science Advances*, 8(34), Article eab6254. <https://doi.org/10.1126/sciadv.abo6254>
- Roozenbeek, J., Van der Linden, S., & Nygren, T. (2020). Prebunking interventions based on "inoculation" theory can reduce susceptibility to misinformation across cultures. *Harvard Kennedy School Misinformation Review*, 1, 1–15. <https://doi.org/10.37016/mr-2020-008>

- Scheibenzuber, C., Hofer, S., & Nistor, N. (2021). Designing for fake news literacy training: A problem-based undergraduate online-course. *Computers in Human Behavior*, 121, Article 106796. <https://doi.org/10.1016/j.chb.2021.106796>
- Schleider, J. L., Mullarkey, M. C., & Chacko, A. (2020). Harnessing wise interventions to advance the potency and reach of youth mental health services. *Clinical Child and Family Psychology Review*, 23, 70–101. <https://doi.org/10.1007/s10567-019-00301-4>.
- Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, 8(4), 350–353. <https://doi.org/10.1037/1040-3590.8.4.350>
- Shaya, F. T., & Gu, A. (2007). Critical assessment of intent-to-treat analyses. *Journal of Medical Economics*, 10, 171–177. <https://doi.org/10.3111/200710171177>
- Shehata, A. (2016). News habits among adolescents: The influence of family communication on adolescents' news media use—evidence from a three-wave panel study. *Mass Communication & Society*, 19, 758–781. <https://doi.org/10.1080/15205436.2016.1199705>
- Shenhav, A., Rand, D. G., & Greene, J. D. (2012). Divine intuition: Cognitive style influences belief in God. *Journal of Experimental Psychology: General*, 141(3), 423–428. <https://doi.org/10.1037/a0025391>
- Shnabel, N., Bar-Anan, Y., Kende, A., Bareket, O., & Lazar, Y. (2016). Help to perpetuate traditional gender roles: Benevolent sexism increases engagement in dependency-oriented cross-gender helping. *Journal of Personality and Social Psychology*, 110(1), 55–75. <https://doi.org/10.1037/pspi0000037>
- Sirlin, N., Epstein, Z., Arechar, A. A., & Rand, D. G. (2021). Digital literacy is associated with more discerning accuracy judgments but not sharing intentions. *Harvard Kennedy School Misinformation Review*, 2(6). <https://doi.org/10.37016/mr-2020-83>.
- Smith, E., & de los Santos, T. (2022). Seeing and doing: Exploring the use of journalist videos and simulations to increase news literacy awareness among high school students. *Media Practice and Education*, 23, 301–314. <https://doi.org/10.1080/25741136.2022.2086960>
- Stephens, N. M., Hamedani, M. G., & Destin, M. (2014). Closing the social-class achievement gap: A difference-education intervention improves first-generation students' academic performance and all students' college transition. *Psychological Science*, 25(4), 943–953. <https://doi.org/10.1177/0956797613518349>
- Tamboer, S. L., Kleemans, M., & Daalmans, S. (2022). 'We are a neeeww generation': Early adolescents' views on news and news literacy. *Journalism*, 23, 806–822. <https://doi.org/10.1177/1464884920924527>
- Tamboer, S. L., Kleemans, M., Molenaar, I., & Bosse, T. (2022). Developing a model of news literacy in early adolescents: A survey study. *Mass Communication & Society*, 1–25. <https://doi.org/10.1080/15205436.2022.2048027>
- Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making*, 11(1), 99–113. <https://doi.org/10.1017/S1930297500007622>
- Van der Linden, S., Roozbeek, J., & Compton, J. (2020). Inoculating against fake news about COVID-19. *Frontiers in Psychology*, 11, 2928. <https://doi.org/10.3389/fpsyg.2020.566790>
- Walton, G. M. (2014). The new science of wise psychological interventions. *Current Directions in Psychological Science*, 23(1), 73–82. <https://doi.org/10.1177/0963721413512856>
- Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science*, 331, 1447–1451. <https://doi.org/10.1126/science.1198364>
- Walton, G. M., Murphy, M. C., Logel, C., Yeager, D. S., Goyer, J. P., Brady, S. T., ... Krol, N. (2023). Where and with whom does a brief social-belonging intervention promote progress in college? *Science*, 380(6644), 499–505. <https://doi.org/10.1126/science.ade4420>
- Walton, G. M., & Wilson, T. D. (2018). Wise interventions: Psychological remedies for social and personal problems. *Psychological Review*, 125(5), 617–655. <https://doi.org/10.1037/rev0000115>
- Walton, G. M., & Yeager, D. S. (2020). Seed and soil: Psychological affordances in contexts help to explain where wise interventions succeed or fail. *Current Directions in Psychological Science*, 29(3), 219–226. <https://doi.org/10.1177/0963721420904453>
- Wineburg, S., Breakstone, J., McGrew, S., & Ortega, T. (2016). *Evaluating information: The cornerstone of civic online reasoning. Evaluating Information Online*. <https://stacks.stanford.edu/file/druid:fv751yt5934/SHEG%20Evaluating%20Information%20Online.pdf>.
- Yeager, D. S., Dahl, R. E., & Dweck, C. S. (2018). Why interventions to influence adolescent behavior often fail but could succeed. *Perspectives on Psychological Science*, 13, 101–122. <https://doi.org/10.1177/1745691617722620>
- Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., ... Dweck, C. S. (2019). A national experiment reveals where a growth mindset improves achievement. *Nature*, 573(7774), 364–369. <https://doi.org/10.1038/s41586-019-1466-y>
- Yeager, D. S., Henderson, M. D., Paunesku, D., Walton, G. M., D'Mello, S., Spitzer, B. J., et al. (2014). Boring but important: A self-transcendent purpose for learning fosters academic self-regulation. *Journal of Personality and Social Psychology*, 107, 559–580. <https://doi.org/10.1037/a0037637>
- Yeager, D. S., Romero, C., Paunesku, D., Hulleman, C. S., Schneider, B., Hinojosa, C., ... Dweck, C. S. (2016). Using design thinking to improve psychological interventions: The case of the growth mindset during the transition to high school. *Journal of Educational Psychology*, 108, 374–391. <https://doi.org/10.1037/edu0000098>
- Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research*, 81, 267–301. <https://doi.org/10.3102/0034654311405999>
- Zerback, T., Töpfl, F., & Knöpfle, M. (2020). The disconcerting potential of online disinformation: Persuasive effects of astroturfing comments and three strategies for inoculation against them. *New Media & Society*, 23, 1080–1098. <https://doi.org/10.1177/1461444820908530>