Long-lasting Effects of a Prosocial Counter-Misinformation Intervention in an Informational Autocracy

Gabor Orosz (✉ gaborosz@gmail.com)
ULR 7369 -URePSSS - Unité de Recherche Pluridisciplinaire Sport Santé Société, Sherpas, Univ. Lille, Univ. Artois, Univ. Littoral Côte d’Opale

Laura Faragó
Institute of Psychology, ELTE Eötvös Loránd University

Benedek Paskuj
University College London

Peter Kreko
Eötvös Loránd University  https://orcid.org/0000-0002-4751-4717

Article

Keywords: Family values, Eastern Europe, Misinformation, Prosocial motivations, Wise Intervention

Posted Date: October 14th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2159993/v1

License: ☑️ This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Conservative voters have difficulties distinguishing fake from real news. In Hungarian representative data (N= 991) we found voters of the reigning populist, conservative party rated misinformation more accurate than real ones independently from the news’ political leaning and content. The question arises: what can psychological science do to make government supporters more motivated and capable in the long run to identify misinformation in this cultural, political, and historical context? Designing scalable misinformation-reducing interventions that have a sustainable effect in the long term is a challenge researchers are struggling to surmount. The present work demonstrates a social psychological randomized controlled trial intervention in which young adults (N= 473) were placed in an expert role and were requested to write a letter to digitally less competent relatives explaining six strategies that can help them to discern misinformation. Compared to the active control group, there was an immediate (d = 0.27) and four weeks later a general long-term effect (d = 0.30) on distinguishing fake news from real ones. This long-term effect was especially salient among participants who support the conservative populist government (d = 0.48). The present work is among the first providing comprehensive quantitative analysis on the gravity of the misinformation problem in an Eastern European country with a pioneering intervention attempt that can be appropriate to a context where misinformation appears regularly in the mainstream media.

Introduction

Conservative voters seem to be making less use of their cognitive capacities while consuming news. A summary of 11 American studies found weaker analytical thinking among political conservatives, compared with liberals, but this was not related to the intellectual capacities of conservatives, rather their motivation to use these cognitive capacities (Deppe et al., 2015; Jost, 2017; Jost et al., 2017). The Hungarian context is ideal to investigate this question. In 2010 the Hungarian right-wing conservative party Fidesz, led by Viktor Orbán, won a super-majority at the parliamentary elections and set about dismantling democratic checks and balances (Bozóki & Hegedűs, 2018), solidifying their grip on power through multiple electoral cycles. Right-wing populist policies were implemented on a wide scale (Boda et al., 2022) and the establishment of media dominance made possible the building of a parallel reality and state-sponsored dissemination of misinformation across the country – turning Hungary into an informational autocracy (Krekó, 2022). Few studies have investigated the psychological ramifications of this media landscape on citizens and to our best knowledge none that tried to intervene to reduce susceptibility to fake news.

Hungary may be the canary in the coal mine. Canaries are more sensitive to toxic gases than humans. If a canary dies, it indicates the imminent danger to miners. Young and fragile democracies such as Hungary can be more sensitive to misinformation campaigns than more robust ones. The situation in Hungary may not be an isolated phenomenon, but an indication of a concerning trend. Keeping this in mind, in the present article, we provide insights about the gravity of the misinformation problem in Hungary (Study 1) and demonstrate a novel intervention approach that builds on the social motivations
of Hungarians, particularly conservative voters, equipping them with the ability to recognize fake news and distinguish them from real ones (Study 2).

**Study 1. The Gravity Of The Misinformation Problem In Hungary**

In Study 1, people rated the credibility of fake and real news that enabled us to compute their ability to discern misinformation and to compare people by political leaning.

**Method**

**Participants**

A reputable polling company—that made the most accurate predictions in the past three elections—gathered a sample representative of the country's population ($N=991$, $M_{age}=50.23$, $SD_{age}=16.07$, range=18-91 years; 51% female; 46.8% high school, 32.5% received a post-secondary degree) who use the Internet at least once a week in April 2021. Among them, 38.6% were supporters of the Orbán government, in line with other representative polls at that time. Respondents were selected randomly from two internet-enabled panels including 520,000 members (more than 5% of the country’s population), and the sample was created using a multiple-step, proportionally stratified, probabilistic sampling method. The study received IRB approval at a Hungarian university in accordance with the Declaration of Helsinki, and with the informed consents of the participants.

**Materials and Procedure**

Following the protocol of Pennycook and Rand (2019), participants received 15 factually fake and 15 real news headlines published on Hungarian fact-checking sites or in mainstream news sources. One third of the headlines were pro-Orbán (ideologically consistent for conservative and populist government supporters [analogous to pro-Trump news], e.g., see Figure 1); one third were anti-Orbán (worldview consistent for supporters of the mainly liberal opposition [analogous to anti-Trump news], e.g., “Péter Szijjártó [Hungarian Minister of Foreign Affairs and Trade]: "It is treason to protest when migrant hordes are besieging Hungary". People were outraged at the Foreign Minister’s words.”), and one third were politically neutral (e.g., “The world is celebrating: a diabetes vaccine has been officially announced. It also helps those already suffering from the disease because it reverses the process.”). Half of the headlines in each category were from fake and half were from real news sites. They were presented as screenshots from a Facebook News Feed including a picture, a headline, and a byline (see Figure 1). The order of the news content was randomly presented to participants and all headlines were pre-tested prior to the data collection (see Supplemental Materials).

The descriptive statistics of fake and real news headlines can be found online in the Supplemental Materials Table S1.

**Analytic Strategy**
Using OLS regressions, we compared fake news discernment scores (accuracy ratings of real news minus fake news) between pro-governmental conservative and oppositional mainly liberal Hungarians.

**Results**

Pro-government, conservative voters had lower average discernment scores (average accuracy ratings of real news minus average accuracy ratings of fake news) than the anti-government, mainly liberal opposition voters, $b=0.52$, $t(989)=8.56$, $p<0.001$, $d=0.54$. This difference did not simply reflect a striking split by partisanship, but also that conservative pro-government respondents rated fake news as more accurate than real news on average - while anti-government, mainly liberal voters rated real news more accurate than fake ones (see Figure 2).

**Discussion**

In the US, Pennycook and Rand (2019) found Democrats significantly better at discerning misinformation from real news than Republicans. In their study however, US residents perceived real news as more credible than fake ones irrespective of political ideology. In Hungary, this was not the case: the average media truth discernment of pro-government voters was negative in all types of news (ideologically matching, mismatching, or politically neutral, see details online in the supplemental materials Figure S1.) meaning that Hungarian pro-governmental voters gave more credit to fake than to real news. Looking at the transformation of the Hungarian media landscape since 2010, this might not shock everyone, but the gravity of the misinformation problem raises questions: What can social psychologists do in this situation? In particular: is there any way to motivate conservative voters to improve their vigilance for misinformation? This question is essential as prior intervention attempts were not as effective for conservative voters as for liberals (see Rathje et al., 2022).

**Study 2. A Prosocial Intervention To Motivate Eastern Europeans To Spot Misinformation**

Various approaches aim to boost people’s resistance to misinformation. Most focus on supporting the individual’s cognitive skills or enhancing their perceptiveness, which lead to immediate or short-term effects that quickly fade. For example, prior ‘nudging’ interventions, that prime people with a simple question: “To the best of your knowledge, is the claim in the above headline accurate?” made readers slow down and motivated them to engage their cognitive capacities in the evaluation of upcoming headlines. Unfortunately, to our best knowledge, these interventions never had long-term effects and their immediate effect is also limited among conservative voters (Rathje et al., 2022). Another approach aims to develop digital skills and competencies, such as digital media literacy programs that try to impart strategies for spotting misinformation. There are only a few studies of this kind that demonstrated measurable effects in media truth discernment after one month. Among these attempts, only a few met the important criteria of scalability (Guess et al., 2020; Roozenbeek et al., 2022). We based our intervention on this work and aimed to modify it in a way that might strengthen the long-term effects in
digitally vulnerable groups, such as pro-Orbán conservatives. Before detailing the psychological mechanisms leveraged in our intervention, we would also like to set the Hungarian socio-cultural context in which the intervention was to work.

Cultural and Historical Context of Family-based Prosocial Motivations

Hungary is not exceptional among Eastern European countries concerning its value structure: family- and security-related values have been identified as central to the population since their first measurement dating back to the 1960s (Andorka, 1992; Beluszky, 2000; Csite, 2009). Family and the home gained special importance after WW2 and into the 1950s when the ruling Soviet Union-backed socialist party banned more than 90% of clubs, unions, and organizations not under direct political control (Hankiss, 1990). In the 1960s and 1970s, people were allowed to make decisions more autonomously in certain parts of their private lives, but any sort of social organization that could potentially challenge the ruling ideology was suppressed or banned. Significant efforts were made to depoliticize society and make people focus on their narrow social groups and compatible family-related or hedonistic values without allowing them to freely organize broader social groups. By the 1980s, the retreat to private life had reached extreme levels. In 1982, 85% of Hungarians said they “would not sacrifice themselves for anything beside their family” - the corresponding figures for Ireland, Denmark and Spain were 55%, 49%, and 38%, respectively (Hankiss, 1990).

Following the fall of the Soviet Union and the transition from socialism to capitalism in 1989, financial security and uncertainty avoidance became more important for Hungarians, but the central role of the family did not diminish (Füstös & Szakolcai, 1999). Between 1978 and 1998, the safety of the family was consistently found to be the first or the second most important value people endorsed. In brief, Hungarian society never recovered from the disintegration of social ties in the socialist era, and various political regimes used this decay for their own benefit instead of systematically investing in the reconstruction of a social fabric that could support collective values (and altruism that transcends the boundaries of the ingroup). The present work demonstrates how these values can still be leveraged to bring about positive societal outcomes that benefit the collective: how the motivation to protect an elderly relative can make the youth more vigilant to spot misinformation.

Prosocial Motives to Identify Misinformation

Can these family-related prosocial values be capitalized on to encourage people’s discernment of misinformation? Prosocial values can provide a powerful motivational force as people can be more motivated to do extra work for the people they care about in contrast to themselves (e.g., Grant, 2007; Grant & Shandell, 2022). For example, health care professionals’ hand-hygiene behavior can be more effective if they are reminded of the positive implications for their patients and not for themselves (Grant & Hofman, 2011). In the field of education, prior studies showed that prosocial motives could help students perform well in monotonous and boring jobs in the US (Paunesku et al., 2015; Reeves et al., 2020; Yeager et al., 2014). In these interventions, students combined self-oriented and self-transcendent (prosocial) goals to be more persistent. Prosocial goals could pertain to close relations (e.g., their family
members, friends) or members of the broader community (e.g., work colleagues, people living in the same town, or everybody on the planet). Considering that Hungarians can be motivated to make an extra cognitive effort for their loved ones, we decided to focus on close relatives and the responsibility to help the digitally vulnerable elderly members of the family.

**Psychological Mechanisms Supporting Prosocial Values**

Unlike most prior misinformation interventions, that put people in a ‘learner’ role (implicitly assuming relative incompetence), our approach granted participants a digital expert role within their families. One of these mechanisms was attributing a stable and positive digital expert role to the participants which derived from their young generational status compared to their elderly family members’ generational status (Miller et al., 1975). The intervention framed youth vigilance to online misinformation as a role modeling behavior for aiding the digitally less competent older generation. Both adolescents and young adults are sensitive to *social-status related signs*, so we aimed not only to highlight their expertise but to demonstrate how digital responsibility can be a source of respect and higher status in their peer groups and beyond (Yeager & Dweck, 2018).

The intervention was framed in a digital learning mindset perspective (similar to Yeager et al., 2019). It means that we framed the digital strategies as competencies that can be developed through (1) making an effort, (2) choosing elaborate learning strategies, and (3) asking for advice from more competent people. Effort is needed as there are barriers that can prevent the spotting of misinformation, such as fatigue or a wandering mind. The core of the intervention material was based on a list of strategies (Guess et al., 2020) that help people to spot misinformation. Participants were encouraged to learn about these strategies through the testimonials of fellow students.

Some of the testimonials used *self-distanced self-talk* to create perspective from the experiencing and analyzing parts of the self (Kross et al., 2014). For example, to demonstrate the “Questioning information that is outrageous” misinformation spotting strategy, participants could read in the excerpt: “*I used to be pretty scared about all sorts of things I read online. For this reason, I have figured out that if a news story comes across with some threatening message, I will stop for a moment and ask myself: John, can this really hurt you now or is it an unfounded nonsense that was written just to make you to be scared?*”. The main reason for suggesting this strategy was not only to encourage the participants to use their analytic thinking and flag the outrageous information, but also to reduce the potentially threatening aspects of misinformation and enable its observation from a safe mental distance.

The testimonials led to a *self-persuasive exercise* similar to the prosocial purpose intervention of Yeager et al. (2014). Writing a letter to their elderly loved ones, youngsters were supposed to indirectly persuade themselves, which can lead to longer-lasting effects than direct persuasion (Aronson, 1999). Explaining the strategies can also induce hypocrisy by highlighting the distance between the advice and their own behavior, which also motivates them to behave in accordance with their advice (Aronson et al., 1991; Stone et al., 1994, see Figure 3, see the simplified timeline of the study in Figure 6).
Conservative Political Orientation as Vulnerability to Misinformation

Though misinformation poses a threat to the whole of society, there are social groups that are especially vulnerable to the harmful effects of misinformation. Research from the US and Western Europe (Baptista et al., 2021) also shows that conservative voters are more likely to be deceived by misinformation (Allcott & Gentzkow, 2017) due to their heightened susceptibility to threatening information (Fessler et al., 2017; Miller et al., 2016) and less deliberative cognitive styles compared to liberals (Jost, 2017). Based on Study 1 and international literature from Western countries (e.g., Allcott & Gentzkow, 2017; Miller et al., 2016), in the Hungarian context, conservative government supporters can be expected to be more vulnerable to misinformation.

Skill-improvement interventions that ultimately aim to change beliefs can be more effective among people who have more room to learn and develop their competencies. For example, growth mindset interventions can lead to more academic performance benefits among students with lower prior grades. Students’ grade point average gains are multiple times larger among at-risk students compared to students with good prior grades (Yeager et al., 2016; Yeager et al., 2019). Following the “more-room-for-development” notion, we expect that pro-governmental (~conservative voters) can benefit more from the present intervention. Our intervention leverages family bonds to motivate young adults to build resistance against misinformation, and these family-values are more important for conservatives (Vogl & Freese, 2020), hence we have reason to expect that they will benefit from our intervention the most.

Fake-news Interventions, Scalability, and Long-term Effects

There are several promising interventions fighting the malicious effects of misinformation – yet, most of them have short-term effects. For instance, accuracy-nudging interventions (Chen et al., 2015; Fazio, 2020; Lutzke et al., 2019; Pennycook & Rand, 2021; Pennycook et al., 2020; 2021; Salovich & Rapp, 2021) are popular and effective as they are short, inexpensive, readily scalable, and autonomy supportive (Pennycook & Rand, 2021), which makes them easy to implement for social media platforms (Kozyreva et al., 2020; Pennycook et al., 2021). Nevertheless, nudges orient users’ behavior in the present moment (e.g., via using the actual situational and environmental cues) without changing their interpretations or narratives on the news consumption situation (Kozyreva et al., 2020). For instance, an accuracy reminder before reading a headline increased truth discernment in participants’ subsequent sharing intentions (Pennycook et al., 2020). However, if these accuracy nudges are not present on social media sites or news outlets, they cannot change people’s behavior. Hence, we argue that nudges remain a double-edged sword if they don’t facilitate a shift in the reasons why spotting misinformation is important; such a shift is more likely to lead to long-term and recursive effects (Walton & Wilson, 2018; Yeager & Walton, 2011).

Even if fake news is among the hottest research topics in psychology, only a handful of studies examined long-term effects of interventions. These studies applied mainly inoculation (Maertens et al., 2020; 2021; Zerback et al., 2020) or other forms of competency-fostering cognitive techniques (e.g., digital media
literacy interventions, Guess et al., 2020). Inoculation-based and competency-fostering cognitive techniques require people's cooperation and in-depth engagement (Hertwig & Grüne-Yanoff, 2017; Kozyreva et al., 2020), therefore they are not readily scalable (Pennycook & Rand, 2021), and sometimes they take several hours, which makes them difficult to implement among the general public. Though more recent attempt translating these approaches to video formats may address these issues (Roozenbeek et al., 2022). Unfortunately, their immunizing effects either vanish completely in the long-term (e.g., effect sizes dropped from $d=0.13$ to $d=0.05$ in India; Guess et al., 2020), or decreased considerably over a short period of time (e.g., effect sizes dropped from $d=0.95$ to $d=0.28$ within one week and from $d=0.20$ to $d=0.08$ in the US; Guess et al., 2020; Maertens et al., 2020). In sum, the efficacy of the scalable misinformation interventions on accuracy or discernment rating decreases or vanishes in one to four weeks. We anticipate more success in the present study by going beyond 'nudging' and competency building and to engage deep-seated motivations that facilitate sustained behavior change over time.

**The Present Study**

Digitally vulnerable people need scalable interventions, which not only change people's behavior in the short term but ideally make a lasting difference. Our first goal was to fill this gap and design a scalable intervention with long-lasting effects that not only trains digital literacy skills but provides good reasons why spotting misinformation is relevant to participants. Our second goal was to design an intervention that can be effective beyond Western and independent cultural contexts where misinformation is peripheral compared to a context where the government controls most media platforms and is the most important source of systematic misinformation (Barlai & Sik, 2017; Juhász & Szicherle, 2017). Finally, in light of the results from Study 1, our third goal was to construct an intervention that can mainly encourage conservative (pro-governmental) voters to spot misinformation.

**Methods**

**Participants**

Participants of the intervention were students enrolled in various majors at a Hungarian public university and 462 young adults reached the randomized intervention or control materials ($M_{\text{age}}=22.24$; $SD_{\text{age}}=4.63$; 72.59% female; 93.77% Caucasian, 33.97% first-gen). They were recruited from a class and, although participation in the study was voluntary, all the students who opened the link agreed to participate. There was some attrition in the follow-up, as data was collected in the middle of the Hungarian fourth wave of the COVID-19 pandemic and some students did not provide an appropriate Student ID that prevented us from matching their follow-up to their intervention data. This led to an attrition of 28.55% of students with intent-to-treat follow-up data from 71.45% of the allocated students ($N=338$, $M_{\text{age}}=22.20$; $SD_{\text{age}}=4.32$; 73.51% female; 95.12% Caucasian, 33.53% first-generation, for summary see Figure 4). This study was conducted with the ethical approval of anonymized university’s institutional review board, in accordance with the Declaration of Helsinki, and with the informed consents of the participants.

**Procedure**
Welcomed and briefed on the study, participants first filled out a measure assessing the frequency of social media use, “bullshit” receptivity (Pennycook et al., 2015), and demographics and then proceeded to their randomly assigned condition.

For the treatment group the exercise was framed as a contribution to an online media literacy program developed for the parents’ and grandparents’ generation. Participants read about six scientifically supported strategies (all adapted from Guess et al., 2020), accompanied by peer testimonials, that could help one to spot misinformation online. Participants were then asked to compose a letter to a close family member that summarized the strategies and to reflect on the best arguments and advice that would persuade their reader to utilize these strategies in life. The strategies included skepticism for headlines; looking beyond fear-mongering; inspecting the source of news; checking the evidence; triangulation; considering if the story is a joke (see an example in Figure 5).

For the control group the exercise was similar in both structure and content but was not related to news. It was framed as a contribution to a social media literacy program developed for the parents’ and grandparents’ generation. It was related to practices of the parents or grandparents that the younger generation finds especially embarrassing. The structure of the control material was very similar; however, the topic of fake news did not appear. Participants read about six practices violating tacit norms of social media use. These practices included using Facebook’s feed instead of private messaging; virtual bouquets for birthdays and name days; inappropriate emoji use; ‘funny’ profile pictures; inadequate device handling during video calls; mass invites for online games. They were then asked to compose a letter to a close family member that summarized the practices and to reflect on the best arguments and advice that would persuade their reader to avoid these practices online.

Measures

Fake news accuracy and media truth discernment: Following the well-established protocol of Pennycook and Rand (2019), we captured perceived news accuracy by having participants rate real and fake news items on a four-point scale. There were eight real and eight fake news items, half of them with political content. The items overlapped with the item set we also used in Study 1, and all had been pre-tested in a prior culturally adjusted replication study of Pennycook and Rand (2019; see Supplemental Materials). Fake news accuracy scores were related to the accuracy rating of the fake news items, whereas the media truth discernment scores were calculated by subtracting the mean perceived accuracy of fake items from the mean perceived accuracy of real items. Further self-reported and behavioral measures were implemented immediately after the intervention and also in the follow-up (descriptive statistics of these variables can be found online in the Supplemental Materials Table S2).

The timeline of the study is shown in Figure 6.

Analytic Strategy
Using OSL regression models, we examined the effect of the condition (treatment vs. control) on fake news accuracy and media truth discernment scores (based on Pennycook & Rand, 2021). We ran this analysis for the immediate and the long-term (one month) results. Our primary analysis was intent-to-treat analysis including all respondents who reached the end of the survey and provided data about their fake news accuracy (see the pre-registration: anonymized). We could not analyze the responses of those students who dropped out before the outcome measures, as we did not have relevant outcome variables (the accuracy of fake- and real news). For assessing the long-term results, we could only use the data of those respondents who were randomly allocated to the treatment and the control groups, and also finished the accuracy ratings in the follow-up. Besides the main effect of the intervention, we were interested in the effect of the intervention among pro-government respondents. As the present work is part of a broader project, we slightly diverted from the original and pre-registered analytic plan. Originally, we aimed to analyze fake news accuracy scores (without considering real news accuracy scores) and not the difference between real and fake news scores (media truth discernment). However, to be able to harmonize with the results of Study 1, here we report media truth discernment and focus on results that are related to political partisanship.

Results

Preliminary Analyses and Attrition

We examined overall attrition (independent of condition) and differential attrition (condition-dependent) in both samples. Considering both overall and differential attrition, students retained, as compared to those not retained, did not differ by sociodemographic (age, gender, residence, parental level of education, first-generation status), social media use (Facebook and Instagram use frequency), or relevant psychological characteristics (“bullshit receptivity”, need for cognition and analytic thinking; \( p > 0.109 \); see Supplemental Material Table S2 for these individual differences). We found that the proportion of minority students was marginally higher in the first round of data gathering compared to the second one (\( z = -2.07; p = 0.038 \)), but there was no condition difference in this case either. Considering the above-mentioned individual differences, there were no baseline differences between the treatment and the control groups (\( p > 0.13 \)).

Primary Analyses

Fake News Accuracy and Discernment in General. Overall, the results showed that the intervention produced significant immediate, \( \beta = -0.15, t(420) = -3.27, p = 0.001, d = 0.31 \), and long-term accuracy rating changes, \( \beta = -0.11, t(333) = -2.26, p = 0.024, d = 0.25 \), relative to the control condition considering the whole sample. It was also true for the immediate, \( \beta = 0.16, t(420) = 2.85, p = 0.005, d = 0.27 \), and long-term, \( \beta = 0.17, t(333) = 2.81, p = 0.005, d = 0.30 \), media truth discernment score (mean real news accuracy scores minus mean fake news accuracy scores, Figure 7, left panel). The intervention’s main effect on fake news accuracy and discernment were significant above and beyond all assessed demographic (age, gender, level of education, place of residence), media use frequency (Facebook, Instagram), and psychological variables (analytic thinking, need for cognition, and bullshit receptivity).
Fake News Discernment among Pro- and Anti-government Voters. Government supporters immediately after the intervention benefitted marginally more than their peers in the control group in terms of distinguishing fake from real news $b=0.22$, $t(102)=0.1982$, $p=0.054$, $d=0.38$. However, this difference became significant and salient one month later in the follow-up, $b=0.27$, $t(85)=2.394$, $p=0.019$, $d=0.48$. Anti-government voters also benefitted from the intervention in short-, $b=0.14$, $t(315)=2.27$, $p=0.024$, $d=0.24$, and long-term, $b=0.16$, $t(238)=2.16$, $p=0.032$, $d=0.27$, as measured by media truth discernment.

Discussion

While these lines are written, we are in the middle of the fifth coronavirus wave and an exploding European energy crisis. Hungary's neighbor, Ukraine, is under the siege of the Russian military, a country with a massive international disinformation arsenal that greatly impacts Eastern and Central European countries (Boksa et al., 2019). In this context, information over misinformation can mean literally life or death for Eastern and Central Europeans. The goal of the present intervention was to leverage family-oriented prosocial motivations that constitute one of the most important value pillars of Hungarian culture (Hankiss, 1990). According to the results, young people—by explaining fake news discerning strategies (e.g., skepticism for headlines or looking beyond fear-mongering) to their elderly family members—developed long-lasting digital distinguish fake from real news. These effects held a month later and were particularly high among conservative, pro-government voters.

Unlike prior fake news interventions using nudges (Chen et al., 2015; Fazio, 2020; Lutzke et al., 2019; Pennycook et al., 2020; 2021; Salovich & Rapp, 2021), inoculation techniques (Banas & Miller, 2013; Basol et al., 2020; Bryanov & Vziatysheva, 2021; Cook et al., 2017; Jolley & Douglas, 2017; Maertens et al., 2020; 2021; Roozenbeek & van der Linden, 2019; Roozenbeek et al., 2020; Scheibenzuber et al., 2021; Van der Linden et al., 2017; 2020; Zerback et al., 2020), or digital competence building (Basol et al., 2020; Banerjee et al., 2017; Guess et al., 2020; Hameleers, 2020; Kahne & Bowyer, 2017; McGrew et al., 2019; Roozenbeek & van der Linden, 2019; Roozenbeek et al., 2020; Scheibenzuber et al., 2021; Van der Linden et al., 2020), the present program addressed why fake news detection is important. It could catalyze recursive psychological processes that keep young adults vigilant to discern fake news and the effect of the intervention can remain over a longer period of time. Prior fake news interventions very rarely found long-term effects (e.g., Guess et al., 2020) and more vulnerable groups rarely benefitted over and above others. Previous studies found only minimal effects of accuracy nudge interventions on conservative voters (Rathje et al., 2022). Our study found a good tool for reducing the susceptibility of conservative voters to fake news. We believe that the main reason for this is that prior interventions focused on the individual and their cognitive motivations but did not use the motivations of relevant social bonds.

Theoretical Contributions

How can psychologically wise interventions (Walton & Wilson, 2018) contribute to misinformation research? Instead of competing among nudges and inoculation techniques or competency-building interventions, how is it possible to combine the best of these approaches to advance the fight against misinformation?
All these approaches put the individual at the heart of the intervention, but it does not mean that they cannot incorporate a framing of interdependence. Nudging interventions (see Pennycook et al., 2020; 2021) might be combined with these techniques to remind people of their expert role and responsibility towards elderly relatives. Accuracy nudging (Chen et al., 2015; Fazio, 2020; Lutzke et al., 2019; Pennycook & Rand, 2021; Pennycook et al., 2020; 2021; Salovich & Rapp, 2021) can be effectively implemented not only to make people temporarily elaborative right after the intervention, but they can remind people and reinforce the narrative and the reasons why it is important to be vigilant to spot misinformation. In inoculation and competency-building interventions (Banas & Miller, 2013; Basol et al., 2020; Banerjee et al., 2017; Cook et al., 2017; Guess et al., 2020; Hameleers, 2020; Jolley & Douglas, 2017; Kahne & Bowyer, 2017; Maertens et al., 2020; 2021; McGrew et al., 2019; Roozenbeek & van der Linden, 2019; Roozenbeek et al., 2020; Scheibenzuber et al., 2021; Van der Linden et al., 2017; 2020; Zerback et al., 2020) emphasizing the expertise of the participants can reduce reactance and can provide an additional layer of motivation and persistence to incorporate the intervention materials (instead of putting them in a more or less inferior position in which it is assumed that they are digitally incompetent) (Hertwig & Grüne-Yanoff, 2017; Kozyreva et al., 2020).

On the other hand, inoculation interventions could add prosocial motives by encouraging to explain to loved ones the reasons why it is important to play the fake news game. However, in these works it might be also essential to consider the cultural and social-class-related values. The Hungarian context provided excellent ground for the harnessing of interdependent, family-based prosocial values (Füstös & Szakolcai, 1999); yet, in other countries (e.g., in the USA) prosocial motivations towards broader communities or more specific identity-relevant social groups (such as ethnic, class, neighborhood-related groups) might be more effective to motivate people to spot misinformation (Reeves et al., 2020, Yeager et al., 2014).

No interventions aiming to shape narratives can influence everyone equally (Yeager & Walton, 2011). In Hungary—resonating with societal values—the protection of families is the dominant political narrative, especially recently. The country has no ministry of education or health care, but there is a minister for families – the former one was inaugurated as the president of the republic earlier in 2022. Based on the massive governmental communication, Hungarian families receive various sorts of psychological threats and also readymade protection from the state. For example, the government is tamping constantly the “threat of migrants” (asylum seekers in 2015), the threat of George Soros, the threat of Brussels, the threat of high utility bills (etc.), and they provide a protection to these threats. Most importantly, the government’s communication resonates very well with the dominant values of the society: the protection of families. For these reasons, we found that there is a salient effect of the intervention among the voters of the pro-Orbán conservative political party.

Applied Contribution

Researchers and policy makers are hoping for interventions with optimal scalability, long-term effects, and solid efficacy on vulnerable groups (Pennycook & Rand, 2021). Yet, immediate implementation to wide audiences is usually difficult as the framing of these interventions is critical. Our
approach was effective among young adults in the context that made it plausible for them to give advice that helps the preparation of an intervention for elderly people. We may achieve similar effects by scaling this program through online adverts and getting many young adults to write a kind letter to their elderly relatives. We can achieve lower attrition, but less involved participants if we hire polling companies who externally reward the participants[1]. A media campaign for promoting the intervention and supported by private companies (with the support of a responsible tech company) and public agencies (with the support of national agencies responsible for life-long and digital learning) can be an approach that can reach masses on both sides of the political spectrum especially in these times when there is strong social media consumption.

The mechanisms the present intervention used were based on prosocial family values. We suspect that threatening situations such as the war neighboring Hungary can increase the perceived threat and the heightened threat can orient people even more towards their loved ones. Therefore, we suspect that the present intervention can have even stronger effects now than in the fall of 2021. Future studies might test this assumption. From our perspective, the value structure of the other Eastern European countries is not very different from the Hungarians (Schwartz & Bardi, 1997). After WWII, the Soviet occupation made significant efforts to cut off extra-familiar bonds not only in Hungary (for a study on Poland see Arato, 1981) as the spontaneously emerging social connections and groups could pose potential danger to the Soviets’ political power. Similarly to Hungary, the consequence of these atomizing policies in terms of anomic wounds have not been healed in these countries (Hankiss, 1990). Hence, the narrow scope of prosocial intentions (focusing on friends and family) holds true in other countries of the region, which suggests that misinformation interventions could successfully capitalize on these familial bonds. Future studies might explore the application of this intervention content in other cultural contexts.

Limitations

Despite being a randomized controlled trial with small and not significant (differential and also overall) attrition and carefully pretested behavioral measures, the present work is not without limitations. As it was among the first steps of the intervention development, our sample was not nationally representative and comprised only university students. We cannot, therefore, extrapolate to how effective this intervention among young adults is in the general population. It was only tested in the Hungarian context; we do not know yet whether the intervention is similarly effective in other Eastern and Central European countries. Finally, similar to almost all interventions, as a dependent variable we did not measure the actual social media behavior over time but used Pennycook and Rand’s (2019) assessment method.

[1] We asked five polling company and none of them could guarantee even a 40% attrition rate independently from the rewarding structure.

Conclusion

Misinformation interventions can save lives in pandemic and wartimes – both of which have recently been witnessed by Central and Eastern Europe, a region that is re-emerging as a political and
informational buffer zone between Russia and the Euro-Atlantic alliance. The core psychological motive underlying our intervention, prosocial family values, was inspired by the theoretical and methodological approach of wise interventions (Walton & Wilson, 2018; Yeager et al., 2014), but also influenced by the communication strategies of the populist Hungarian government that often builds on similar values to rally its electorate (e.g., when they frame asylum seekers as threats to Hungarian families). The present intervention, framed around the protection of vulnerable family members, demonstrates that these values can be leveraged to boost critical thinking and misinformation discernment and could be highly effective in cultures where interdependence and family values are important. This intervention might be only a first step towards a new generation of misinformation interventions that combines prior individuum-focused strategies with social ones.

References


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


784-793. https://doi.org/10.1177/0956797615571017


**Declarations**

**Funding and acknowledgement:** 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. The data gathering of Study 1 was supported by the funds of Jigsaw. The first author was supported by Northern French Strategic Dialogue (Phase 1 & 2) and STaRS Grants. Special thanks to Williams Nuytens and Pasquale Mammeone for their institutional support. Special thanks to the Schmidt foundation postdoc grant that made possible to study prosocial wise interventions at Stanford University. The second author was supported by ÚNKP-22-4 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund, Grant/Award Number: ÚNKP-22-4. The fourth author was supported by the János Bolyai Research Fellowship of the Hungarian Academy of Sciences (grant nr: BO/00686/20/2) and by the ÚNKP-20-5 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund (nr: ELTE/12912/1(2021). Special thanks for the insightful comments to Rocky Cole, Dezso Nemeth, Manual Galvan, Greg Walton, Hugh Willan, Lilla Török, and Juan Ospina.

**Author contributions statement**: G.O., L.F., and B.P. designed the study, L.F. and G.O collected samples. G.O. analyzed the data and produced the figures. G.O., L.F., B.P. and P.K. provided interpretation of the results. All authors participated in writing up the manuscript.

**Data availability**

Study 1: The dataset that supports the findings is openly available on Open Science Framework (DOI 10.17605/OSF.IO/VCF36) and can be found here: https://osf.io/vcf36/

Study 2: The dataset that supports the findings will be openly available on Open Science Framework after publication and it is submitted with the manuscript along with the R script.

**Additional information**

We declare that the authors have no competing interests as defined by Nature Human Behaviour, or other interests that might be perceived to influence the results and/or discussion reported in this paper.
Figures

One example of a pro-Orbán fake news headline

Notes. English translation of the headline: “We [Hungary] will be among the best 5 countries after the migration crisis has subsided” Participants were asked about is the accuracy of the claim in the headline (not at all accurate/not very accurate/somewhat accurate/very accurate; based on Pennycook & Rand, 2019).
Figure 2

Accuracy rating of real news minus fake news along partisanship on a representative sample

Notes. Based on a representative sample, conservative Hungarian pro-government respondents rated fake news as more accurate than real news on average (left side); however, anti-government, mainly liberal voters rated real news more accurate than fake ones (right side). Error bars represent standard errors. The
y-axis represents standardized discernment scores in terms of real news minus fake news accuracy ratings. Higher scores indicate better discernment between fake and real news. Scores below zero show more perceived accuracy in fake than real news; meaning that they gave more credit to fake news than to real news.

Figure 3

Psychological Mechanisms of the Self-persuasive Message
Figure 4

Sample characteristics and attrition.

Original strategy:

**Find more articles.** If another news source does not report the same story, this may indicate that the news is false. If the story is reported by more reliable sources, it is more likely to be true.

Adam's opinion:

“I used to be enthusiastic about sharing shocking news so that all my acquaintances would know first from me what was going on in the world. Then one day one of my older waterpolo training buddies, who I looked up to very much, told me he saw my post and I should delete it immediately. He said that if you share new and interesting news, they will think you are really informed. If, on the other hand, you share fake news without reason, they will look at you like a dumb kid, even though I don’t think you are. I went back to the pool and was constantly thinking about when I could delete my post. I’ve been sharing a lot of things since then, but before that, I always check several sites to see if it’s really true.”

Adam, a 24-year-old economist

Figure 5

*A Testimonial Explaining One of the Six Strategies.*
Figure 6

Timeline.

Notes. Participants of the intervention benefitted more than their peers in the control group in terms of distinguishing fake from real news. It was true if we considered the full sample (left), anti-government voters (middle), and the difference between the control group and the treatment was striking among government supporter conservative voters (right). Error bars represent standard errors. The y-axis represents discernment score in terms of real news minus fake news accuracy ratings, y axis covers approximately 2 SDs. Higher scores indicate better distinction between fake and real news.

Figure 7

Fake News Discernment among Pro- and Anti-government Voters One month after the Intervention

Notes. Participants of the intervention benefitted more than their peers in the control group in terms of distinguishing fake from real news. It was true if we considered the full sample (left), anti-government voters (middle), and the difference between the control group and the treatment was striking among government supporter conservative voters (right). Error bars represent standard errors. The y-axis represents discernment score in terms of real news minus fake news accuracy ratings, y axis covers approximately 2 SDs. Higher scores indicate better distinction between fake and real news.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.
• prosocfnanalysismarkdown.txt
• SupplementalMaterial.docx