Can We Shield Citizens Against Misinformation Through Digital Literacy Training and Fact-checks?

Antonella Bandiera,[1] Horacio Larreguy,[2] John Marshall,[3]
Daniela Pinto Veizaga,[4] Manuel Quintero[5]

[1] Department of Political Science, ITAM
[2] Departments of Economics and Political Science, ITAM
[3] Department of Political Science, Columbia University
[4] Data Science Program and Department of Economics, ITAM
[5] Departments of Applied Mathematics and Economics, ITAM

We partnered with a Bolivian fact-checker organization, Chequea Bolivia, to address the misinformation problem in the Global South by providing digital literacy training and fact-checks. We recruited two samples: one received the training material through an online workshop (Online) and the second through WhatsApp messages (WhatsApp). Participants were randomly assigned to only receive fact-checks (Verification), only training (Course), both training and fact checks (Course and Verification), or nothing (Control). In general, we found that both delivery methods successfully inoculate participants. The Online method has a large effect on knowledge and builds more capacity to discern fake news, thus continuing to consume the same amount of news from social media. In turn, the WhatsApp delivery method increased citizens' awareness of misinformation, ultimately altering participants' level of consumption. Last, we found that those that received verifications through WhatsApp messages share less fake news.

Introduction

Misinformation has been an increasing concern around the globe because it might lead people to take adverse actions or decisions based on false premises. Misinformation extends to any relevant subject in society, from health and economic issues to politics. More recently, misinformation has become pervasive around the COVID-19 pandemic. Misinformation's growth and consequences are more severe in the Global South, where social media - the primary source of misinformation - is rapidly extending among populations with low levels of digital literacy. For instance, misperceptions about COVID-19 (*1*) and other diseases like Zika are widely believed (*2*) in the Global South.

Most misinformation studies focus on debunking (exposure to fact-checks) to correct misinformation in the short term. Debunking interventions can be effective. (*3*), show that corrections from qualified fact-checkers effectively reduce misperceptions. Also, corrections from trusted institutions are more successful among those with substantial initial misperceptions (*4*). However, debunking can have backfiring effects; for instance, exposure to general warnings can decrease beliefs under truthful information (*5*). Thus, the evidence on correcting misinformation through fact-checks, specifically to change individuals' beliefs when priors are strong, is mixed (*6*). Moreover, when an intervention successfully changes beliefs, these shifts tend to be ephemeral or fail to be persistent over time (*7, 8*).

Another technique to correct misinformation is "prebunking," or *inoculation* by exposing people to past misinformation, warnings, and teaching tactics. Inoculation can engender resistance to future encounters with misinformation (*9, 10*). Also, effectively warning individuals of the threat of misinformation can reduce the extent to which it affects them (*11*). Several studies show the benefits of *inoculation* as a tool to spot and build resistance against misinformation (*12–14*).

A third technique to avoid misinformation is digital literacy training. However, in-person and online digital media literacy interventions are scarce and might have backfiring effects by making individuals over-confident in their abilities to discern fact from fiction (*15*). Moreover, literacy training had shown null or low effectiveness. In particular, in-person literacy training significantly impacts those more educated and has no effect in less literate environments (*16*). Additionally, short literacy training by exposure, correction, and revealing the importance of misinformation did not affect participants' ability to identify information in a low literacy context (*17*).

Most of these interventions on debunking, inoculation, or digital literacy training focus primarily on the Global North, with few on the Global South. Bowles et al. (*18*) showed that a long-term intervention disseminated via WhatsApp in the Global South can have a sustained effect in altering participants' ability to discern false information, increase their knowledge of verification, and reduce their trust in social media. However, it fails to shift participants' consumption behaviors. Furthermore, Bowles, Larreguy, and Liu (*1*) showed that debunking through WhatsApp messages via a reliable source alters participants' behavior in the short term and induces people to reevaluate their COVID-19 misconceptions in Zimbabwe. Although the COVID-19 pandemic sparked interest in field experiments in the Global South, there is still a need for further research to understand the effectiveness of different intervention approaches in the Global South to shield citizens against misinformation and evaluate lasting effects on varied outcomes to find potentially scalable solutions to the problem of misinformation.

We partnered with Chequea Bolivia, a recent but renowned fact-checker organization that, since June of 2019, has been uninterruptedly debunking information, sharing corrections and tactics on how to discern truthful from false information, as well as delivering digital literacy courses to lessen the consequences of misinformation in Bolivia.

Through Chequea Bolivia's Facebook and webpage, we recruited two samples of Bolivian

individuals. The first sample aimed to recruit subscribers for an online workshop on misinformation (Online sample). The second sample aimed to attract WhatsApp users to acquire similar information through WhatsApp messages (WhatsApp sample). Both samples were randomly assigned to either receive all the information from the course after completion of the endline survey (Control group) or one of our three treatment arms. The first treated group had a chance to take both the training – through an online course or similar content delivered through WhatsApp messages – and the fact-checks published on Chequea Bolivia's online site during the intervention period (Course and Verification). The second group only took the online course or obtained the course information through WhatsApp messages (Course). The remaining participants were assigned to receive only the fact checks published on Chequea Bolivias' webpage (Verification).

This study was designed in this way for several purposes. First and more importantly, we wanted to combine previous methods of assessing information: literacy training, inoculation, and fact-checks on a single intervention. This study differs from preceding short-term literacy training interventions (*17*) in that ours is longer. The online workshop lasted roughly a month and the WhatsApp sample received course information for 10 days to maximize attention. Second, we aim to compare, given the same information, which, or for what purposes, each delivery method – Online vs. WhatsApp – is more effective at internalization and provoking changes in broader behaviors such as consumption. For instance, the online course was self-paced, respondents were only given a deadline to complete the course's units, and they had two in-person practice sessions. On the other hand, WhatsApp users only got messaged with content from the course once a day. Third, we want to assess which treatment is more effective: receiving the course information, receiving only the verifications, or getting both.

To this end, we first analyze how participants interacted with the course information or verifications from Chequea Bolivia. Following prior studies, we then study whether combining

digital literacy training and fact-checks can have an effective impact on reducing misinformation (*19*) or whether news literacy messages on social media correct misinformation (*20*). Thus, we measure whether the intervention increased their interest and attention to misinformation and their ability to discern between truthful and false sources of fake news. Third, we assessed the change in trust in different information sources and participants' increase in self-reported and indirect knowledge to discern and verify the information. Last, we evaluate whether tools to identify misinformation and shifts in trust of information sources can extend to shifts in broader behaviors (*21, 22*).

In conclusion, we aim to combine debunking, inoculation, and literacy training through a course delivered either online or via WhatsApp messages and the dissemination of fact-checks. Moreover, our study focuses on the Global South and its long-term effects rather than on the Global North and its short-term outcomes. Furthermore, we seek to establish a comparison between two delivery methods and understand which one is more effective: the online course or WhatsApp messages. Lastly, we seek to understand if this intervention can increase individuals' awareness of misinformation and shift their consumption, sharing, and verifying behaviors.

Sample recruitment, treatment assignment, and content distribution

During the last week of January, Chequea Bolivia launched an open call to sign-up for a free virtual workshop called *Habilidades críticas y dígitales contra la desinformación* (Critical and digital skills against misinformation, in English) with the possibility to earn an official certificate after completion. Through its webpage and social media accounts, Chequea Bolivia recruited participants to take the workshop. The Facebook post title was: *Atención Bolivia, ¿Quieres ser parte de la lucha contra la desinformación?* (Attention Bolivia: Do you want to be part of the fight against misinformation? in English) and featured information about the

virtual workshop and photographs depicting the topics covered: over-information, journalism, critical thinking, assessing information, digital media, digital content, digital disinformation, misinformation, and fake news, digital tools, and verification methodology. Afterwards, we incentivized participants to answer a baseline survey of approximately 15 minutes with the chance of being one of three USD 150 prize winners. These participants make up the Online sample. Participants of the WhatsApp sample were recruited similarly through social media posts from Chequea Bolivia.

Respondents from both samples were invited to subscribe to a WhatsApp list from fact-checker Chequea Bolivia and add the number to their phone contacts to be able to participate in the study. Participants first indicated their consent to be part of the project and agreed to answer the baseline surveys; then, they were screened for eligibility.

We first verified that prospects from both samples had sent a message to subscribe Chequea Bolivia's WhatsApp number and add it to their contacts. Next, individuals were only eligible for any intervention if they were at least 18 years old, lived in Bolivia at the time of the intervention, had a Bolivian cellphone number, and a WhatsApp account (if applicable). Then, if participants met the eligibility criteria, they could answer the rest of the questions from the baseline survey. At the end of the baseline survey, respondents from both samples were invited to provide their Facebook and Twitter accounts to measure treatment effects in their posts before and after the intervention. Additionally, the WhatsApp sample was required to subscribe to a second WhatsApp list from the research team and add the number to their contacts to receive future communications.

In the end, 310 people completed the baseline survey of the online workshop and 528 the WhatsApp survey, see Table 1. Both the survey for the Online and the WhatsApp sample reported participants' demographics, social media use, information consumption, interest in misinformation, opinions on the likelihood of false of different information sources, knowledge

to identify and verify misinformation, and were asked to report attitudes on hypothetical scenarios regarding receiving fake and truthful information and whether they would trust, verify, and share such news. Lastly, the baseline survey included questions on broader behaviors on consumption, sharing, and verifying information.

Participants that completed the baseline survey were randomly assigned, for each sample separately, to either receive no course information until March (Control), or to any of the three treated groups: Course and Verification, Course, and Verification. We block-randomized respondents for each sample separately by demographic characteristics, exposure to misinformation, knowledge to discern and verify misinformation, and participants' beliefs of how likely are different news sources to provide false information. Table 1 summarizes the number of members in each treatment arm for each sample at baseline. After the treatment groups were formed, participants either took the course or received only fact-checks, both, or nothing. The course is aimed at literacy training to give individuals the knowledge to navigate their information environment mindfully and critically (*23*) and to stimulate critical thinking by understanding the negative impacts of misinformation (*24*), as well as corrections through the distribution of fact-checks.

The online course started on February 3rd and continued for over a month; it had four main units and an additional practical workshop. The units were divided into (1) misinformation and digital citizenship; (2) analyzing news through critical thinking; (3) the digital ecosystem; and (4) fact-checking and digital abilities. Each unit required roughly two hours to complete. In addition, the content of each element is very interactive, given that it contains videos, podcasts, lectures, and quizzes designed particularly for this course. Furthermore, each unit had a final summary to reinforce the information learned. The first section aimed to recognize the importance of media literacy in constructing a more digitally informed society in a high-misinformation environment. The second unit served to develop critical thinking abilities to

evaluate information. The third module sought to teach and identify the most relevant features of the digital world. The fourth element taught Chequea Bolivia's methods and usage of digital tools as support to verify misinformation. Lastly, the course ended with an assessment that included practical applications to reinforce each unit and a final test.

On the other hand, the WhatsApp course ran from February 17 to 27, with the material divided into three parts: (1) Introduction to misinformation, (2) Critical thinking and biases, and (3) Fact-checking. The first module ran for five days and included a 5 minute, 33 seconds video on methodologies to verify fake news, infographics on fake images, videos, fake audios, and deep fakes. The second module ran for three days consisting of a 5 minute, 25 seconds video on the importance and use of critical thinking to interpret, analyze, evaluate, infer, and explain information, followed by an infographic on information consumption and biases in information processing. Last, the third and fact-checking module lasted three days and included a 5 minute, 33 seconds video emphasizing the means to identify, verify, and rectify false information; and two infographics summarizing what citizens can do to detect misinformation. Appendix Table S1 contains links and details on the WhatsApp course contents.

During March, we asked participants to respond to a ten-minute endline survey. Table 1 shows 205 and 258 responses for the Course and WhatsApp endline sample, respectively. Participants in the control group received the course information following the survey. All respondents were incentivized with the opportunity to participate in a lottery for three prices of USD 150 each, and with the final certificate after completing the survey.

Table 1: Response rates by treatment assignment and by sample

Sample	Treatment	Baseline	Endline	Response rate
Sample Course delivered online	Control	83	66	0.80
	Course and Verification	76	45	0.59
	Course	73	35	0.48
	Verification	78	59	0.76
	Total	310	205	0.66
Sample Course delivered via WhatsApp	Control	132	72	0.55
	Course and Verification	132	65	0.49
	Course	131	66	0.50
	Verification	133	55	0.41
	Total	528	258	0.49

Through these surveys, we first analyze the effectiveness of the delivery methods as measured by the interactions with the course information or fact-checks from Chequea Bolivia. Then, we analyze whether combining digital literacy training and fact-checks can effectively reduce misinformation by studying an extensive range of outcomes.

Sample demographics

Our combined endline sample from both surveys with 463 observations is demographically balanced across treatment arms; see Table S6. All respondents are above 18 years of age, reside in Bolivia at the time of the surveys, and have access to the internet. When comparing our sample with national and representative surveys of the Bolivian population, we encounter relevant differences among samples. We compared our sample with the latest Americas Barometer survey[1] (2021) with $3,002$ observations and the most recent Latin Barometer survey[2] (2020) with $20,077$ responses. Figure S6 presents the results for comparing age, gender, and education variables among our experimental sample, the Americas Barometer, and the Latin Barometer overall. We further restricted both national samples to respondents that reported having access to the internet in their household either by a paid contract or just by receiving an internet signal.

[1]Find data in `https://www.vanderbilt.edu/lapop/bolivia.php`.
[2]Find data in `https://www.latinobarometro.org/latContents.jsp`.

Relative to the unrestricted Americas Barometer and Latin Barometer, respondents in our experimental sample are younger. In general, younger people are easier to reach on social media, (sub-figure S6a), less likely to be female (sub-figure S6c), and more educated (sub-figure S6e). Although, the difference between the endline survey and both restricted samples to internet users is narrower, the discrepancies are still significant (sub-figure S6b, S6d, and S6f). See also Appendix Table S23 for more details. We later address these differences in our sample and find that our results will likely expand beyond our sample to the broader internet user population in Bolivia.

Results

The endline survey of the online course sample had 205 responses, and the WhatsApp endline sample had 258 observations showing an overall response rate of 0.66 and 0.49, respectively. Both samples exhibit significant differential attrition, see Panel A of Table S4. However, the WhatsApp sample presents less differential attrition. To account for this differential attrition, we implement the Inverse Propensity Score Weighting (IPSW) technique to estimate treatment effects accounting for the difference in attrition, see the Supplementary Materials for additional details on our IPSW estimation. Notice that Panel B of Table S4 shows less differential attrition after implementing IPSW for both samples.

We presented our main findings using standardized Index Covariance Weighting (ICW) summary indexes of the outcome families at endline and created the analogous index at baseline when feasible. see Appendix Table S12 for details on the variables and directions of each index's component. Each index's component is set in the direction that we hope the treatment should lead to an increase in the outcome.

Table S6 shows that both samples are demographically balanced across treatment groups. However, Panel A of Table S5 shows imbalances for the main indexes at baseline for the Online

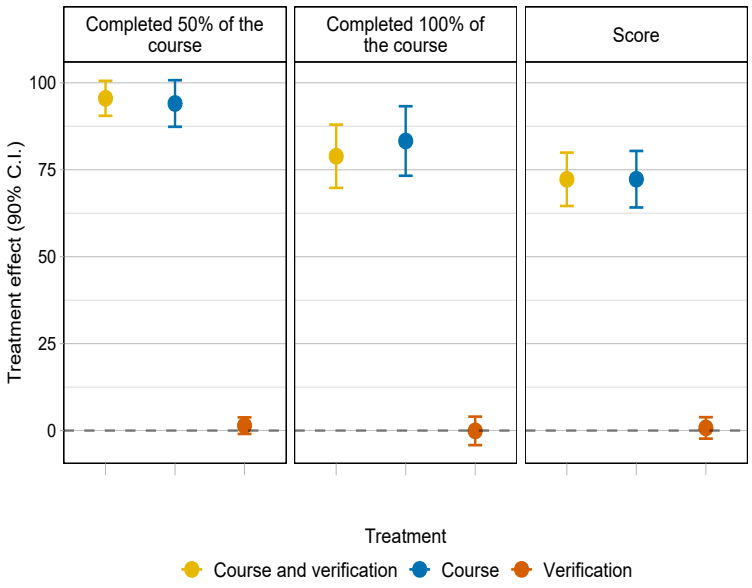
sample and a fair balance at baseline for the WhatsApp sample; tables S7 to S11 show in more detail balance for all baseline covariates. Noteworthy, the WhatsApp sample is significantly more balanced than the Online sample across most baseline variables. These imbalances are somewhat addressed when computing the IPSW (see Supplementary Materials for more details) resulting in a more balanced sample, see Panel B of Table S5. To conclude that we have a more balanced sample across treatments, we compared the number of coefficients statistically significant at least at the 90% level. Notice that the balance for the Online sample in Panel A.1 of Table S5 has 12 coefficients with $p < 0.1$ and Panel B.1 has only 2 such coefficients out of 33; a major improvement in balance. Panel A.2 has 4, and Panel B.2 only 5 coefficients such that $p < 0.1$ out of 33.

Throughout the analysis, we report treatment effects for both the Online and the WhatsApp samples on endline outcomes variables. Across all effects, we compare the different treatment arms, Course and Verification, Course, and Verification, relative to the Control group. For comparison and robustness purposes of the main effects, we implement three estimation techniques with the OLS method and adjust for randomization block fixed effects. Estimations further include the lagged dependent variable (if measured at baseline). First, the main results are simple OLS estimations including all covariates in the outcome family at baseline. Second, we estimate a Weighted Least Square (WLS) model with IPSW weights to account for differences in attrition. Third, we estimate an OLS model with covariates selected by a LASSO model from all baseline variables and indexes to increase statistical power and account for the lack of balance. HC_0 robust standard errors are used throughout.

Using administrative data from Chequea Bolivia, we first show that the information was successfully disseminated for both samples. Figure 1 shows that for the online endline sample 95.5% ($p < 0.01$) of the Course and Verification group and 94% ($p < 0.01$) of the Course group completed at least 50% of the assessments in the course, and 78.9% and 83.3% ($p < 0.01$) of

each treatment completed 100% of the assignments, respectively. Furthermore, both treatment groups, on average, scored above $70/105$ ($p < 0.01$) in the survey questions, which implies that more respondents answered correctly the tasks related to the information delivered to them. As expected, participants assigned to the Verification treatment arm did not take the course (see analogous results for the baseline sample and more details in Appendix Table S2).

Figure 1: Percentage of respondents that completed 50% and 100% of the course, and the achieved score for the sample whose course was delivered online

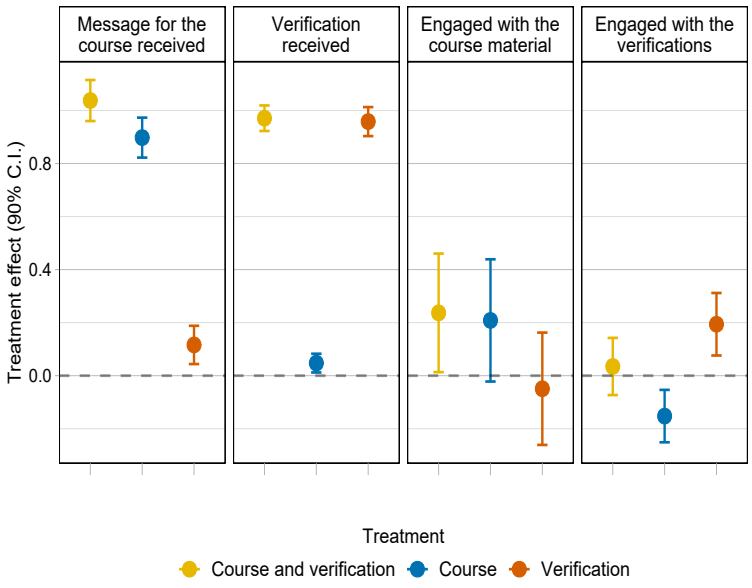


Notes: Columns (1) and (2) are indicators for completing the corresponding percentage of the course material. Column (3) is the final score after completion of the assessment section. The estimates and 90% confidence intervals in each box are from separate OLS regressions, where the outcome is the corresponding label regressed on treatment indicators and randomization block fixed effects. The outcomes are in Panel B of Table S2.

Similarly, Chequea Bolivia's data shows that the information was successfully delivered to the WhatsApp sample. The left panel of Figure 2 exhibits that both the Course and Verification and the Course treatment groups successfully received the messages related to the course material ($p < 0.01$). Notably, the Course and Verification group displays higher reception and interaction (93%). As expected, the Verification treatment shows no difference from the

Control group. Likewise, the second panel of Figure 2 shows that the Course and Verification and the Verification arms received fact-checks ($p < 0.01$). The third and fourth panel of the same Figure presents the results of active interactions with the course material and verifications sent.[3] An active interaction is an individual responding or commenting to any course message or verification sent at any time during the intervention. Similar to the messages received, the third panel shows that the Course and Verification and Verification treatment groups actively engaged with the course material ($p < 0.01$). The right panel shows that although the Course and Verification group did not show an active response to the verifications sent, the Verification group actively responded to the fact-checks ($p < 0.05$). Additional details for all panels can be found in Appendix Table S3.

Figure 2: Interaction with videos, messages, and verifications for the sample whose course was delivered by WhatsApp



Notes: Each column is an indicator variable for the corresponding outcome. The estimates and 90% confidence intervals in each box are from separate OLS regressions, where the outcome is the corresponding label regressed on treatment indicators and randomization block fixed effects. The outcomes are in Panel B of Table S3.

[3] Although Chequea Bolivia did not keep records of the engagement with fact-checks for the entire sample, this data is available for the WhatsApp sample.

After assessing the treatment's effective dissemination, we analyze the treatment effects on the main endline indexes relative to the Control group. Appendix Table S13 summarizes all the results explained below and Appendix Tables S14 to S22 contain the treatment effects for the components of each index separately.

Misinformation importance

We first measure the respondents' perceptions of misinformation after the intervention by asking how likely it is that misinformation contributes to different problems in Bolivia. Problems ranged from health issues to electoral concerns, including ideological polarization, violence, and hatred towards certain groups of people in Bolivia, see Appendix Table S14 for details on the problems asked about. The left panel of Figure 3a shows statistically insignificant effects across all treatment arms for the Online sample on the ICW index of indicators when misinformation affects Bolivian problems The Course group of the WhatsApp sample presents a marginally significant decrease ($p < 0.1$) when relating misinformation to Bolivian problems, see panel 1 of Figure 3b. However, we found that the data supported the null model over the alternative model with a Bayes Factor of 0.248 for this marginally significant decrease, see Table S28. The data also supports the null model for the rest of coefficients with a Bayes Factor of 0.106 - 0.818. The IPCW and LASSO estimates are consistent with these findings, see panel 1 of Appendix Figure S2 or column 1 of Table S13 for additional details.

Likelihood of false sources and Distrust sources

Next, we analyze participants' beliefs about how likely different news sources are to distribute misinformation (likelihood of false) and how much they trust the information in these sources. We examine the treatment effects on both likelihoods of false and distrust by dividing each category into two indexes, capturing traditional or social media sources. Traditional sources

include radio / TV, newspaper, and the internet, whereas social media include WhatsApp, social media in general, and casual conversations with family and friends.

First, Appendix Figure S1 supports the null model over the alternative, with Bayes factors of 0.119 - 0.743, see Table S28, in testing the decrease in the likelihood of false of traditional sources for both the Online and the WhatsApp survey. Mirroring this behavior, we statistically rejected effects on trusting more traditional sources with Bayes factors ranging from 0.074 to 0.516 for both samples. Notably, Appendix Figure S2 displays a decrease in the likelihood of false of traditional sources for the Course and Verification group of the Online sample for the LASSO estimations ($p < 0.05$); with a consistent statistical insignificant effect for the rest of the treatments in both samples. The IPSW and the LASSO estimates capture no effect in trusting more traditional sources.

On the other hand, panel 2 of Figure 3b shows an increase in the belief of falsehood of information received on social media across all treated groups of the WhatsApp sample in similar magnitudes (0.29 - 0.36 SD increase), relative to the Control group ($p < 0.05$, $p < 0.1$, and $p < 0.05$ for Course and Verification, Course, and Verification respectively). When implementing IPSW and LASSO, the effects are consistent in magnitudes and significance.

Consistent with the increase in the falsehood of news coming from social media, participants from the WhatsApp sample also increased their distrust of information coming from social media. In particular, from panel 3 of Figure 3b, we notice that the Course group and the Verification one distrust social media more than the Control group in similar magnitudes (0.34 and 0.23 SD increase; $p < 0.05$ and $p < 0.1$, respectively). Both treatment effects are validated through the LASSO estimates, see columns 3 and 5 of panel C.2 of Table S13, with an increase in magnitude (0.39, and 0.27, respectively) and an increase in significance to $p < 0.01$ for the Course group. Also, the IPSW method is robust in magnitude and significance, see Appendix Figure S3.

The treatment effects of the WhatsApp sample display larger effects and are consistent with the belief that the intervention should reduce their trust in social media and increase the likelihood of false news coming from social media sources.

Attention to misinformation

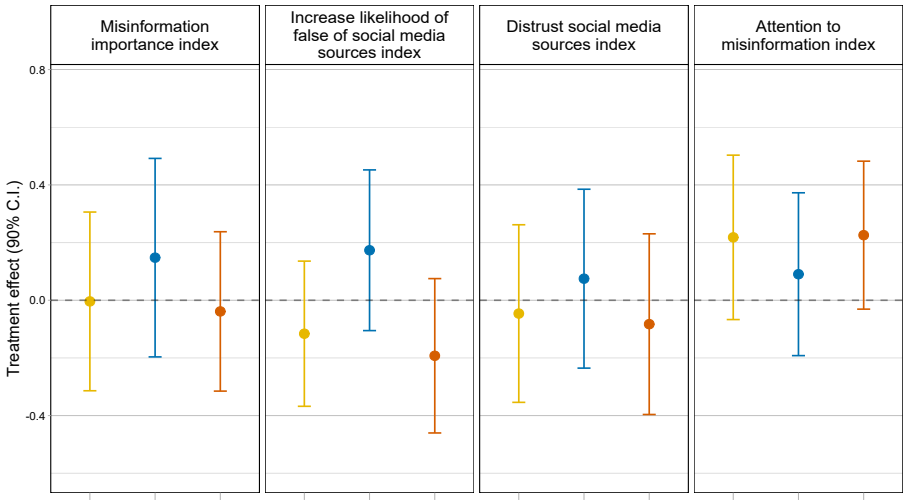
Then, we measure the attention paid to misinformation by respondents from both surveys by asking how much attention they pay to misinformation when looking at the news.

The results in panel 4 of Figure 3a show that the Online intervention did not increase participants' attention to misinformation at the 90% level. However, estimates show that those that receive verifications via WhatsApp (Course and Verification and Verification) have similar and large positive effects (0.22 and 0.23 SD increase; $p = 0.21$ and $p = 0.15$, respectively). Although marginally insignificant, the data support the alternative model over the null with a Bayes factor of 2.049 for the Course and Verification group, see Table S28. These effects hold with IPSW and LASSO estimates, with the Verification group significance increasing above the 90% level, see panel 2 of Appendix Figure S2a.

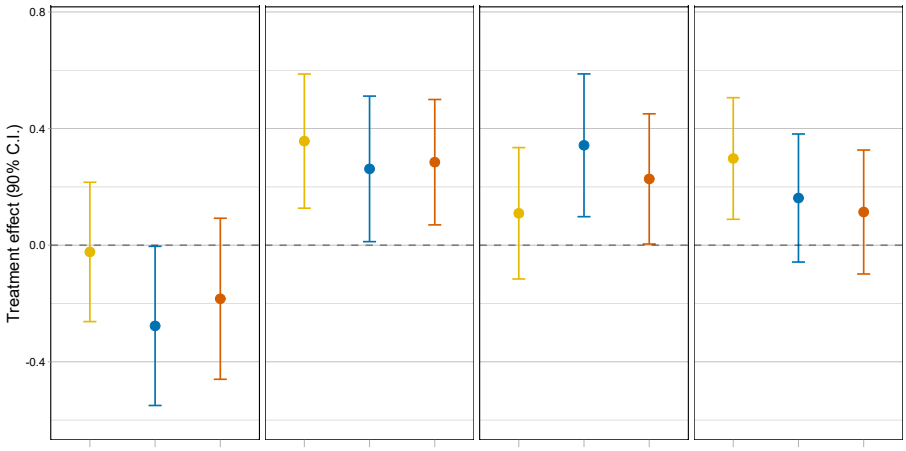
Regarding the WhatsApp sample, panel 4 of Figure 3b shows an increase in the attention paid by the Course and Verification treatment group (0.30 SD increase; $p < 0.05$), which is also supported by the corresponding IPSW and LASSO coefficients at the same significance level and magnitude in panel 2 of Appendix Figure S2b. Furthermore, panel 4 of Figure 3b shows that the Course and the Verification groups have a positive and large effect (0.16 and 0.11; $p = 0.23$ and $p = 0.38$, respectively). While underpowered, everything delivered via WhatsApp (the course for the WhatsApp sample and verifications for both samples) has large magnitudes.

Figure 3: Treatment effects on mechanisms: misinformation importance, misinformation attention, likelihood of false of traditional and social media sources, and distrust of traditional and social media sources

(a) Sample course delivery online



(b) Sample course delivery WhatsApp



Notes: The estimates and 90% confidence intervals in each box are from separate OLS regressions. The labels are the corresponding dependent variables regressed on treatment indicators and randomization block fixed effects. The outcomes are in columns (1), (3), (5), and (6) of Panel A of Appendix Table S13.

Knowledge to identify and verify fake news

We then assess whether participants were able to identify between potentially false and truthful information and if they could verify the veracity of news by different means. We measured how much participants could discriminate false news using self-reported and objective measures. Firstly, participants were asked to report how much knowledge they have to identify whether a piece of news is false or not; secondly, participants were asked to identify characteristics of false news from a list that included correct and incorrect answers, recent misinformation cases in Bolivia, and if they could identify the main sources through which fake news are spread.

Although the left panel of Figure 4 shows uniformly statistically insignificant effects for all treatments in both samples relative to their Control groups at the 90% level, those that took the course online (Course and Verification and Course of the Online sample) have large and indistinguishable estimates (0.20 and 0.25; $p = 0.26$ and $p = 0.21$, respectively). The left panel of Appendix Figure S4a shows that these results hold for the IPSW estimates and improve for the LASSO coefficients (0.21 and 0.37 SD increase; $p = 0.19$ and $p < 0.05$ for the Course and Verification and Course groups respectively).

Respondents' increase in knowledge to verify the information was also measured by asking how much knowledge they possessed, what ways to verify the news they knew, and what fact-checkers they could use to verify the veracity of some information. The right panel of Figure 4a shows that similarly to the reported knowledge to identify information, participants that took the online course increased their knowledge to verify information (0.61 and 0.57 for the Course and Verification and Course, respectively, with $p < 0.01$ for both groups). Likewise, the verification group reported no effect. Panel 2 of Appendix Figure S4a shows that these results are robust in the same significance level for both the IPSW and LASSO estimates.

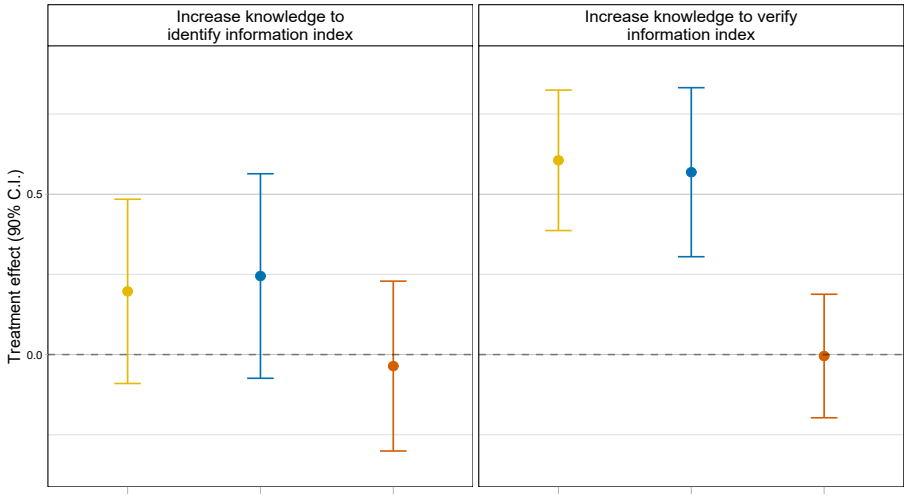
The Course group of the WhatsApp sample also had a marginally significant increase in their knowledge to verify information (0.22 SD increase; $p < 0.1$), and this result is robust to

the IPSW and LASSO results. While underpowered, the Course and Verification group shows a similar effect in magnitude (0.15 SD increase; $p = 0.23$), mainly supported with the LASSO estimate (0.19 SD increase; $p = 0.11$), see Appendix Figure S4b. The data strongly supports the alternative model over the null hypothesis for the Course and Verification group with a Bayes factor of 584.261.

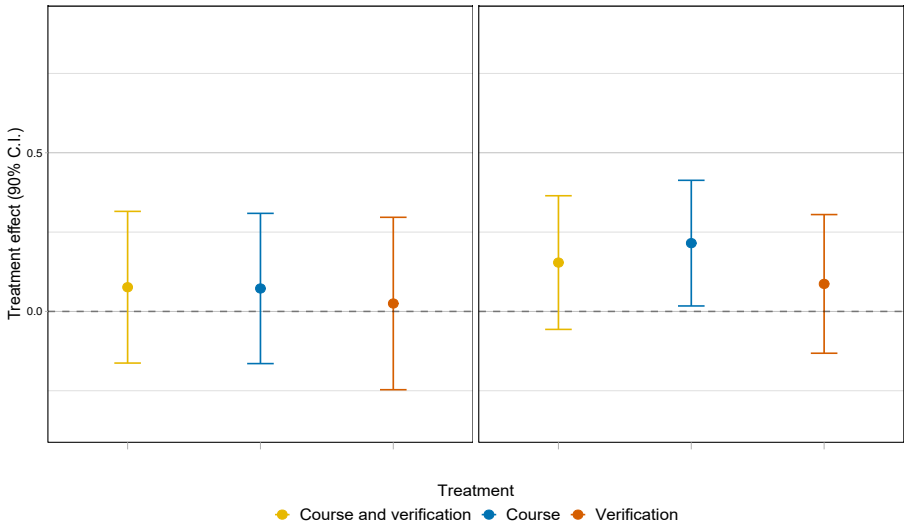
Clearly, the subset of the Online sample that took the course increased their ability to verify information by employing fact-checkers and identify the main characteristics of both fake and truthful news, as supported by Appendix Table S19. Contrary to the prior results, the Online delivery method was more efficient when teaching how to identify and verify information than the WhatsApp intervention.

Figure 4: Treatment effects on knowledge to identify and verify misinformation

(a) Sample course delivery online



(b) Sample course delivery WhatsApp



Notes: The estimates and 90% confidence intervals in each box are from separate OLS regressions. The labels are the corresponding dependent variables regressed on treatment indicators and randomization block fixed effects. The outcomes are in columns (7) and (8) of Panel A of Appendix Table S13.

Information consumption, sharing, and verification

Finally, we turn to evaluate if the shift in participants´ beliefs of how likely news' is to be false depending on their source, their (mis)trust of different sources, as well as the changes in the capacity to identify and verify likely misinformation, affected different behaviors relating to the consumption of news and misinformation. In particular, the intervention measures how each treatment group modified their behavior regarding the consumption of traditional and social media sources, sharing news on social media, and verifying information.

Consumption behavior:

To assess the effect on consumption, we created a separate index for the consumption of traditional media and a second one for social media. As expected and consistent with prior findings in the decrease in the likelihood of false and increase in trust in conventional sources, the right panel of Appendix Figure S1 shows no statistically significant effect in the consumption of traditional sources across all treatment arms for both samples. Similarly, Appendix Table S28 shows that the data support the null model over the alternative with Bayes factors ranging from 0.08 to 0.329. These results are robust to the LASSO estimation and IPSW, see panel 1 of Appendix Figure S5.

Similarly, the left panel of Figure 5a shows consistent results in participants' updated beliefs in the likelihood of false of social media, change in trust towards information from social media sources, and knowledge to identify and verify the information. The Online sample had no statistically significant effect on those that only took the online course and an increase in consumption of social media sources for those that received verifications in similar magnitudes (0.23 and 0.26 SD decrease; $p = 0.17$ and $p < 0.1$ for the Course and Verification and the Verification groups respectively). Both the IPSW and LASSO estimates are consistent with these results, see panel 2 Appendix Figure S5b. Intuitively, the Online sample is more prepared and

confident at identifying and verifying false information, so they continue consuming news from social media.

In turn, the WhatsApp sample presents less consumption of social media sources across all treatment arms (0.30 - 0.49 SD increase; $p < 0.05$, $p < 0.01$, and $p < 0.01$ for the Course and Verification, Course, and Verification groups respectively); see the left panel of Figure 5b. Also, panel 2 of Appendix Figure S5b supports these effects at the same significance level for the IPSW and the LASSO estimates. Intuitively, all treatment groups reported an increase in their belief that social media tend to contain false news and increased their distrust towards information coming from social media. Thus, altering their behavior towards the consumption of news from these sources by consuming less information.

Sharing behavior:

Moreover, panel 2 of Figure 5a also shows the shift in participants' willingness to share information on social media. While the only statistical significant result at the 90% level is the Verification group of the Online sample (0.37 SD increase; $p < 0.05$), the course group also has a large coefficient (0.29 SD increase; $p = 0.21$), which the IPSW estimate improves in magnitude and significance (0.37 SD increase; $p < 0.1$). The LASSO estimate is consistent in magnitude and significance for both groups. Intuitively, participants of the Online sample tend to share less fake news because they know how to identify and verify misinformation.

Although there is a statistical insignificant treatment effect at the 90% level across all arms for the WhatsApp sample, the Verification group shows a large effect (0.24 SD increase; $p = 0.15$) that turns marginally significant ($p < 0.1$) with the IPSW and LASSO estimations, see panel 3 of Appendix Figure S5b. Furthermore, we found that the data support the alternative model over the null with a Bayes factor of 2.067 for the Verification group, see Appendix Table S28.

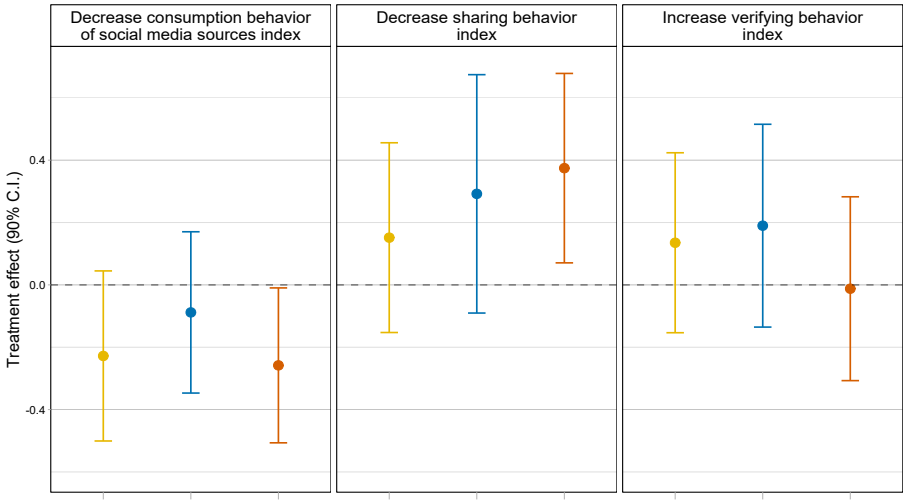
Verifying behavior:

Finally, we asked respondents to self-report how often they verify news that they doubt may be false. The right panel of Figure 5 shows statistically insignificant effects for all treatments in both samples. The IPSW and the LASSO estimates also show no effects. Similarly, we statistically rejected all effects when using Bayes factor with values ranging from 0.071 to 0.283, see Appendix Table S28.

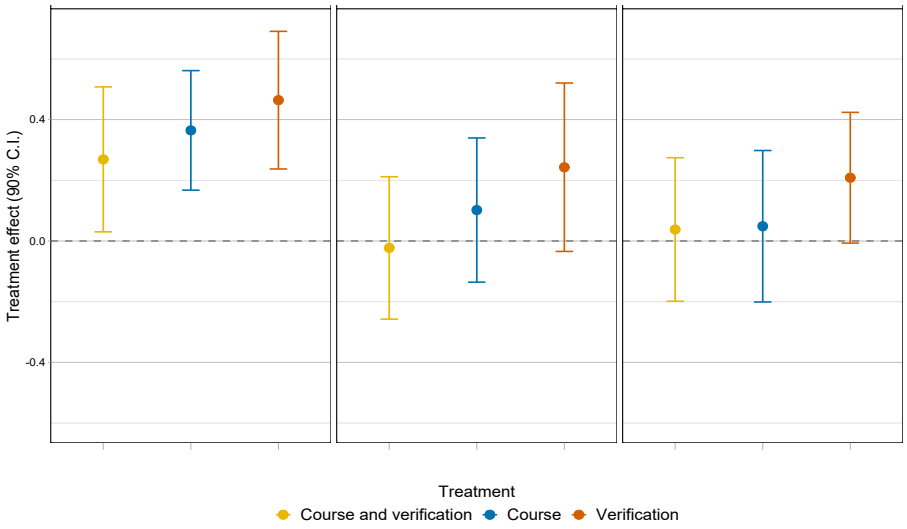
To sum up, the change in consumption of social media sources for the WhatsApp sample suggests that participants' increase in the likelihood of false information that circulates in social media, as well as their change in trust towards social media as sources of information, extend to significant effects in the sources of information from which they consume news. Furthermore, the increase in actual knowledge to identify and verify information from those in the Online sample extrapolates to sharing less misinformation on social media. Additionally, the IPSW and LASSO estimates show that those that received fact-checks from Chequea Bolivia through WhatsApp messages decreased sharing of false news on social media by large magnitudes. Nevertheless, both interventions fail to shift participants' verifying behaviors.

Figure 5: Treatment effects on consumption, sharing, and verifying behavior

(a) Sample course delivery online



(b) Sample course delivery WhatsApp



Notes: The estimates and 90% confidence intervals in each box are from separate OLS regressions. The labels are the corresponding dependent variables regressed on treatment indicators and randomization block fixed effects. The outcomes are in columns (10) - (12) of Panel A of Appendix Table S13.

Robustness checks

As mentioned earlier, Figure S6 shows comparable demographic variables among our experimental sample, the Americas Barometer from the year 2021, and the Latin Barometer from the year 2020 and displays a significant difference between our sample and both national samples. The experimental sample is younger, more educated, and less likely to be female. However, when adjusting with post-stratification weights for these differences, we find that our results are likely to expand beyond the experimental sample to the broader internet user population in Bolivia. Appendix Figures S7 to S10 compare the main family ICW indexes estimations for our sample, which was adjusted to mimic the Americas Barometer and the Latin Barometer independently. Appendix Table S24 weights for the Americas Barometer population and Appendix Table S25 weights for the Latin Barometer population. Both adjustments suggest that our findings largely hold.

Last, we scrapped participants' posts on Facebook and Twitter to check if they shared fake news and fact-checks. We created fake-news models using the neural network model for language processing, BETO, to identify whether a post was fake or not. Then, we made a single ICW index out of the dummy variables from each model to increase statistical power.

First, Appendix Figure S26 shows that out of the 838 baseline responses, 414 reported at least their Facebook or Twitter account. Also, out of the 414, 133 individuals posted at least once during the intervention period. Although we have a small sample, it is well-balanced. Column 1 shows a balanced sample for reporting social media accounts, column 2 presents a balance for posting at least once during the intervention period, and column 3 on the number of posts.

Furthermore, Appendix Figure S27 shows consistent results with the decrease in sharing misinformation from the endline survey data. The Verification group from the Online sample has a similar effect in magnitude and significance (0.39 SD decrease; $p < 0.05$) than the corre-

sponding effect shown in the survey data. Likewise, those that took the course also shared less misinformation (0.39 SD decrease; $p < 0.1$) with a larger coefficient than the treatment effects from the endline survey. Consistently with prior findings, the WhatsApp sample shows a null effect across all treatment arms in decreasing sharing of misinformation. Also, the Verification group of the WhatsApp sample shows the largest effect of the three treatment arms (0.16 SD decrease).

Finally, no treatment arm from either sample presents an effect on sharing fact-checks.

Discussion

The digital literacy training and fact-checking intervention to shield citizens against misinformation led to the following conclusions. First, this work shows that it is feasible to attract individuals to consume literacy training and fact-checking regarding misinformation through digital means.

Second, both delivery methods had different but remarkable results. On the one hand, the more elaborated information delivered through the online course and the longer intervention led the Online sample to get inoculated by teaching them actual tools to identify and verify information. Moreover, those on the Online sample that received fact-checks through WhatsApp messages pay more attention to misinformation. As a result of these effects, participants continued to consume information and did not reduce their consumption behavior; however, because of their increase in knowledge and awareness of misinformation, they tend to share less misinformation on social media.

On the other hand and although the WhatsApp sample received less elaborate information during a shorter period of time, participants' awareness increased more than those that received the course online. The WhatsApp sample had more significant effects on most outcome indexes. It yields positive treatment effects in participants' ability to discern more likely sources

of misinformation. Thus, they increase their distrust of social media as a reliable source of information. Similar to the Online sample, those that received the course and fact-checks through WhatsApp messages increased their attention to misinformation. Moreover, the increase in distrust and attention paid to misinformation led participants that took the course (Course and Verification and Course groups) to increase their knowledge to verify the news. All these combined effects ultimately extend to participants' change in consumption of information from social media sources. Intuitively, individuals are not only more aware of misinformation, but crowds-out consumption of news from social media because they distrust it more. Last, like the Online sample, those that received fact-check messages tend to share less misinformation.

Finally, similar to prior studies but with greater effectiveness, digital literacy training seems to have a positive effect on increasing participants' ability to identify fake news, as supported by both the Online and the WhatsApp samples. Similarly, misinformation corrections disseminated through WhatsApp messages via a trusted source are effective in getting individuals to reassess their misconceptions and correct related behavior within a few days after receiving the messages (*1*). Newly, one of our contributions shows that combining debunking, inoculation, and literacy training in a single intervention through WhatsApp messages seems to alter participants' consumption and sharing behaviors.

Future work should focus primarily on increasing the sample size to support these results. As well as on means to extend these findings to other groups, such as less educated individuals or older citizens. Because the Latin American region relies heavily on the use of television to access news, future interventions should also focus on delivering a misinformation course through television, as in Fotini et al. (*25*) or other means to include such population groups.

References

1. L. H. Bowles, Jeremy, S. Liu, *PLOS ONE* **15**, 1 (2020).

2. C. et al., *Science Advances* **6**, eaaw7449 (2020).

3. F. et al., *Political Communication* **32**, 127 (2015).

4. V. Emily, L. Bode, *Political Communication* **37**, 136.

5. B. et al., *Political Behavior* p. 1073–1095.

6. J. C. R. L. H. Walter, Nathan, Y. Morag, *Political Communication* **37**, 350–375.

7. C. et al., *Psychological Science* **28**, 1531–1546.

8. B. Nyhan, *Proceedings of the National Academy of Sciences* **118**.

9. J. Cook, *Inoculation Theory* pp. 220–236.

10. V. der Linden et al. .

11. K. et al. .

12. M. et al .

13. J. Roozenbeek, S. van der Linden, *Palgrave Communications* **5**, 1–16.

14. S. v. d. L. Roozenbeek, Jon, T. Nygren, *Harvard Kennedy School (HKS) Misinformation Review* .

15. E. K. V. Melissa Tully, L. Bode, *Mass Communication and Society* **23**, 22 (2020).

16. G. et al., *Proceedings of the National Academy of Sciences* **117**, 15536–15545.

17. S. Badrinathan, *American Political Science Review* **115**, 1325–1341.

18. B. et al .

19. M. Hameleers, *Information, Communication Society* **25**, 110–126.

20. L. B. Emily K., M. Tully, *Communication Research* **2**.

21. G. P. Mosleh, Mohsen, D. G. Rand, *PloS One* **215**, e0228882.

22. P. et al. .

23. T. et al. .

24. K. e. a. Clayton, *Political Behavior* **42**, 1073–1095.

25. H. L. Fotini Christia, E. Parker-Magyar, *TSE Working Paper* **1294**.

Acknowledgments

Supplementary Materials

Data and code availability

The replication file with all the data and codes developed by myself using the programming language `R` in `RStudio` software are available in the following Github repository: `https://github.com/Manuquinteroc/MisinformationBolivia`. This repository includes all the raw data from the surveys, data from social media scraping, post-stratification data, data from Chequea Bolivia, all the codes used for data-wrangling, and statistical analysis to generate all the tables and figures presented in this work.

Missing Responses

Though non-responses were rare, these were treated as the median category of the questions asked. For instance, when responding to the question of how often they question whether a piece of news is false, "Don't know" responses were coded as "sometimes". Similarly, "Don't know" responses were coded as the median value for questions related to likelihood, falsehood, and trust questions (e.g. as "neither likely nor unlikely, "neither false nor true", and "neither trust nor distrust").

Block Randomization

To avoid bias and imbalances among treatment conditions, we block-randomized individuals, for each sample separately, by demographics, exposure to misinformation, knowledge to discern and verify misinformation, as well as participants' position towards the likelihood of false of different news sources. We used the `R` package `blockTools` to minimize the multivariate Mahalanobis distance, equation **??**, step by step with an Optimal Greedy Algorithm.

Empirical Specification

The main results presented in this work are from Ordinary Least Square regressions from a Fixed Effect Model. The treatment effects on the main indexes and various outcomes in the Appendix are estimated relative to the Control group using the following specification:

$$Y_{ib} = \alpha_b + \alpha Y_{ib}^{pre} + \beta \mathbf{X}_{ib}^{pre} + \tau_1 \mathbf{T}_{ib} + \tau_2 \mathbf{T}_{ib} \times 1_{\text{WhatsApp}} + \varepsilon_{ib}, \tag{S1}$$

where Y_{ib} is an outcome of interest for respondent i from block b, α_b are the randomization block fixed effects of both samples, and Y_{ib}^{pre} is the analog baseline outcome (if available). $\mathbf{T}_{ib}$ is the vector of treatment assignments (Course and verification, Course, and Verification), $\mathbf{X}_{ib}^{pre}$ is the vector of baseline covariates in the outcome family, or the variables selected by LASSO from all baseline variables[4], and 1_{WhatsApp} is an indicator for responses from the sample course delivered via WhatsApp. The parameter of interest τ_1 captures the average treatment effect for each treatment group in the Online sample and $\tau_1 + \tau_2$ captures the treatment effect for the WhatsApp sample.

Heteroskedasticity-consistent (HC) standard errors are used throughout. We make use of the HC_0 standard errors (White, 1980) that accounts for heteroskedasticity using the fitted residuals ($\hat{u_i^2}$) from estimating equation S1 with OLS:

$$\Sigma = diag\{\hat{u_i^2}\}.$$

Throughout, two-sided t-tests of statistical significance are used to evaluate whether the null hypothesis can be rejected, i.e. the effect of the intervention in the treatment groups is no different than the effect on the Control group.

[4]See the **??**.

Supplementary text

Table S1: Details on the content of the course per day and modules

February	Module / Information delivered	Link
	First module: Introduction to misinformation	
17	Video	Click here
18	Infographic on fake images	Content
19	infographic on fake videos	Content
20	Infographic on fake audios	Content
21	Infographic on Deep Fakes	Content
	Second module: Critical thinking and biases	
22	Video	Content
23	Infographic on information consumption	Content
24	Infographic on most common biases	Content
	Third module: Fact-Checking	
25	Video	Content
26	Infographic on characteristics of fake news	Content
27	Infographic to identify news	Content

Treatment take-up

Table S2: Summary statistics of Online course

	Completed 50 of the course	Completed 100 of the course	Score
	(1)	(2)	(3)
Panel A: Responded to Baseline sample			
Course and verification	76.388***	64.478***	58.069***
	(4.919)	(5.201)	(4.502)
Course	70.930***	56.791***	50.245***
	(5.174)	(5.784)	(4.642)
Verification	2.171	0.749	1.169
	(1.834)	(2.089)	(1.743)
Control mean	0	0	0
Outcome range	[0, 100]	[0, 100]	[0, 102.78]
Observations	310	310	310
R^2	0.635	0.501	0.551
Panel B: Responded to Endline sample			
Course and verification	95.542***	78.889***	72.264***
	(3.036)	(5.499)	(4.640)
Course	94.060***	83.277***	72.285***
	(4.050)	(6.044)	(4.913)
Verification	1.415	−0.088	0.765
	(1.440)	(2.476)	(1.863)
Control mean	0	0	0
Outcome range	[0, 100]	[0, 100]	[0, 102.78]
Observations	205	205	205
R^2	0.933	0.778	0.813

Notes: We report estimates from OLS regression including randomization block fixed effects. Robust standard errors are in parentheses. * denotes p<0.1, ** denotes p<0.05, and *** denotes p<0.01.

Table S3: Respondents that received and interacted with the course's videos/messages and verifications

	Message for the course received	Verification received	Engaged with the course material	Engaged with verifications
	(1)	(2)	(3)	(4)
Panel A: Responded to Baseline sample				
Course and verification	0.872***	0.958***	0.248***	0.093
	(0.062)	(0.022)	(0.080)	(0.063)
Course	0.772***	0.067***	0.238***	−0.041
	(0.073)	(0.016)	(0.079)	(0.055)
Verification	−0.020	0.975***	0.013	0.198***
	(0.059)	(0.021)	(0.070)	(0.064)
Outcome mean	0.558	0.6	0.153	0.125
Outcome std. dev.	0.497	0.491	0.36	0.331
Observations	360	360	360	360
R^2	0.849	0.905	0.264	0.241
Panel B: Responded to Endline sample				
Course and verification	0.925***	0.987***	0.424***	0.059
	(0.068)	(0.026)	(0.124)	(0.097)
Course	0.785***	0.064***	0.396***	−0.128
	(0.086)	(0.022)	(0.127)	(0.086)
Verification	0.004	0.974***	0.138	0.218**
	(0.065)	(0.028)	(0.108)	(0.089)
Outcome mean	0.602	0.591	0.183	0.167
Outcome std. dev.	0.491	0.493	0.388	0.374
Observations	186	186	186	186
R^2	0.873	0.926	0.297	0.356

Notes: We report estimates from OLS regression including randomization block fixed effects. Robust standard errors are in parentheses. * denotes p<0.1, ** denotes p<0.05, and *** denotes p<0.01.

Table S4: Differential attrition

Panel A: Differential Attrition (OLS)	
	Differential Attrition height
A.1 Sample course Delivery Online	
Course and verification	-0.195^{***}
	(0.070)
Course	-0.334^{***}
	(0.072)
Verification	-0.028
	(0.069)
A.2 Sample Course Delivery WhatsApp	
Course and verification	-0.053
	(0.059)
Course	-0.043
	(0.059)
Verification	-0.131^{**}
	(0.058)
R^2	0.267
Panel B: Differential Attrition (IPSW)	
Panel B.1: Sample course Delivery Online	
Course and verification	-0.159^{**}
	(0.077)
Course	-0.314^{***}
	(0.077)
Verification	-0.001
	(0.077)
B.2 Sample Course Delivery WhatsApp	
Course and verification	-0.000
	(0.058)
Course	0.009
	(0.058)
Verification	-0.093^{*}
	(0.054)
Control mean	0.642
Control std. dev.	0.481
Outcome range	[0,1]
Observations	838
R^2	0.279

Notes: We report estimates from OLS regression including randomization block fixed effects. Robust standard errors are in parentheses. * denotes p<0.1, ** denotes p<0.05, and *** denotes p<0.01.

Balance

Table S5: Balance summary on main indexes with and without IPSW

Panel A: Balance (OLS)

	Misinformation importance index	Increase likelihood of false of social media sources index	Trust traditional sources index	Distrust social media sources index	Attention to misinformation index	Increase knowledge to identify information index	Increase knowledge to verify information index	Increase consumption behavior of traditional sources index	Decrease consumption behavior of social media sources index	Decrease sharing behavior index	Increase verifying behavior index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
A.1 Sample course Delivery Online											
Course and verification	−0.079	−0.147	−0.496***	0.186	−0.126	−0.322*	−0.048	−0.135	0.308*	−0.049	−0.208
	(0.179)	(0.189)	(0.173)	(0.190)	(0.144)	(0.175)	(0.172)	(0.183)	(0.168)	(0.181)	(0.177)
Course	−0.505**	−0.377*	−0.387**	0.455**	0.055	−0.529**	−0.341*	−0.072	0.244	−0.181	−0.315
	(0.251)	(0.227)	(0.196)	(0.205)	(0.184)	(0.214)	(0.187)	(0.171)	(0.223)	(0.201)	(0.253)
Verification	−0.231	−0.351*	0.082	0.278*	−0.228*	−0.146	−0.172	−0.083	−0.059	0.079	−0.234
	(0.196)	(0.196)	(0.148)	(0.165)	(0.135)	(0.168)	(0.159)	(0.166)	(0.156)	(0.173)	(0.161)
A.2 Sample Course Delivery WhatsApp											
Course and verification	−0.162	−0.125	−0.351**	−0.023	−0.028	0.026	0.140	−0.298**	0.219*	−0.090	0.162
	(0.147)	(0.131)	(0.177)	(0.166)	(0.082)	(0.140)	(0.116)	(0.120)	(0.122)	(0.131)	(0.120)
Course	0.093	−0.189	−0.190	−0.183	−0.075	−0.030	0.017	−0.311**	−0.015	−0.190	−0.097
	(0.120)	(0.131)	(0.180)	(0.169)	(0.092)	(0.140)	(0.138)	(0.122)	(0.117)	(0.132)	(0.138)
Verification	0.114	−0.066	−0.047	−0.089	0.038	0.390***	0.075	−0.263**	−0.132	0.016	−0.041
	(0.143)	(0.134)	(0.194)	(0.172)	(0.100)	(0.145)	(0.144)	(0.132)	(0.109)	(0.134)	(0.125)
Control mean	0.099	0.102	0.12	-0.059	3.928	0.035	0.002	0.128	-0.11	-0.02	0.029
Control std. dev.	0.82	0.958	0.931	1.016	0.751	1.02	0.993	0.986	0.994	1.017	0.996
Outcome range	[-4,0.8]	[-2.8,1.2]	[-2.6,2.1]	[-3,1.8]	[1,5]	[-3.3,2.4]	[-3.1,2.5]	[-3.2,1.8]	[-1.4,3.8]	[-3,2.1]	[-3.4,1.1]
Observations	463	463	463	463	463	463	463	463	463	463	463
R^2	0.302	0.366	0.302	0.345	0.473	0.408	0.503	0.510	0.519	0.461	0.443

Panel B: Balance (IPSW)

Panel B.1: Sample course Delivery Online											
Course and verification	−0.029	−0.196	0.039	−0.106	0.021	0.083	0.038	0.080	0.014	0.073	0.144
	(0.150)	(0.163)	(0.162)	(0.167)	(0.142)	(0.156)	(0.142)	(0.153)	(0.131)	(0.153)	(0.184)
Course	−0.242	−0.328**	−0.224	0.110	0.106	0.165	−0.149	0.118	−0.140	0.093	0.155
	(0.171)	(0.153)	(0.153)	(0.161)	(0.135)	(0.155)	(0.140)	(0.142)	(0.141)	(0.149)	(0.176)
Verification	−0.247	−0.276*	0.206	−0.168	−0.169	0.120	−0.061	0.164	−0.219	0.129	−0.112
	(0.181)	(0.162)	(0.155)	(0.168)	(0.141)	(0.157)	(0.146)	(0.150)	(0.138)	(0.154)	(0.176)
B.2 Sample Course Delivery WhatsApp											
Course and verification	−0.083	−0.003	0.015	0.025	0.047	0.063	0.307***	0.145	0.087	−0.112	0.109
	(0.119)	(0.111)	(0.120)	(0.119)	(0.088)	(0.118)	(0.108)	(0.103)	(0.105)	(0.112)	(0.113)
Course	0.038	0.000	0.158	0.002	−0.069	0.114	0.187*	0.205**	0.151	−0.257**	0.057
	(0.120)	(0.112)	(0.121)	(0.119)	(0.089)	(0.119)	(0.109)	(0.104)	(0.105)	(0.113)	(0.113)
Verification	0.037	−0.083	0.038	0.150	−0.062	0.074	0.232**	0.141	0.088	0.031	−0.059
	(0.111)	(0.103)	(0.112)	(0.111)	(0.082)	(0.110)	(0.101)	(0.096)	(0.098)	(0.105)	(0.105)
Control mean	-0.069	-0.074	0.046	-0.07	3.916	-0.11	-0.052	0.158	-0.19	-0.026	0.002
Control std. dev.	1.091	0.992	0.971	0.941	0.787	0.972	0.935	0.938	0.884	0.933	0.953
Outcome range	[-3.8,0.6]	[-2.7,1.3]	[-2.5,2.1]	[-3,1.8]	[1,5]	[-3.4,2.5]	[-3,2.6]	[-2.9,1.7]	[-1.4,3.8]	[-3.5,2.1]	[-3.3,1.1]
Observations	838	838	838	838	838	838	838	838	838	838	838
R^2	0.248	0.358	0.242	0.232	0.315	0.278	0.373	0.468	0.412	0.324	0.344

Notes: We report estimates from OLS regression including randomization block fixed effects. Robust standard errors are in parentheses. * denotes p<0.1, ** denotes p<0.05, and *** denotes p<0.01.

Table S6: Balance on demographic variables

	Age	Gender (female)	Level of education
	(1)	(2)	(3)
A.1 Sample course Delivery Online			
Course and verification	−1.269	0.053	−0.318*
	(0.902)	(0.044)	(0.187)
Course	−1.039	−0.017	−0.129
	(0.877)	(0.059)	(0.217)
Verification	−0.426	0.080*	−0.163
	(0.748)	(0.048)	(0.174)
A.2 Sample Course Delivery WhatsApp			
Course and verification	−1.629	−0.006	−0.310
	(1.579)	(0.038)	(0.234)
Course	1.314	−0.003	−0.203
	(1.610)	(0.039)	(0.238)
Verification	0.795	−0.003	−0.002
	(1.702)	(0.041)	(0.252)
Control mean	30.594	0.529	8.833
Control std. dev.	9.761	0.501	1.396
Outcome range	[18,76]	[0,1]	[2,11]
Observations	463	463	463
R^2	0.587	0.898	0.528

Notes: We report estimates from OLS regression including randomization block fixed effects. Robust standard errors are in parentheses. * denotes p<0.1, ** denotes p<0.05, and *** denotes p<0.01.

Table S7: Balance on mechanisms: Misinformation relevance and falsehood

	Misinformation importance index of (1,1,1,1 1,1,1)	Decisions that may affect health	Election of candidates who do not represent the interests of citizens	To falsely discredit or glorify people	Ideological polarization	Violence towards certain people or groups in society	Increase hatred towards certain people or groups in society	Increase likelihood of false of social media sources index of (1,1)	Believe scenario 1 is fake†	Believe scenario 3 is fake††
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A.1 Sample course Delivery Online										
Course and verification	−0.079	0.036	−0.323*	−0.045	0.070	0.007	0.023	−0.147	−0.185	−0.110
	(0.179)	(0.204)	(0.192)	(0.184)	(0.163)	(0.182)	(0.211)	(0.189)	(0.235)	(0.208)
Course	−0.505**	−0.402	−0.361	−0.416*	−0.510**	−0.461**	−0.300	−0.377*	−0.370	−0.382
	(0.251)	(0.271)	(0.227)	(0.247)	(0.256)	(0.235)	(0.237)	(0.227)	(0.254)	(0.271)
Verification	−0.231	−0.155	−0.095	−0.322*	−0.208	−0.337*	−0.303	−0.351*	−0.251	−0.449**
	(0.196)	(0.205)	(0.180)	(0.188)	(0.194)	(0.185)	(0.198)	(0.196)	(0.232)	(0.214)
A.2 Sample Course Delivery WhatsApp										
Course and verification	−0.162	−0.028	−0.241	−0.164	−0.212	−0.055	−0.111	−0.125	0.020	−0.267
	(0.201)	(0.209)	(0.201)	(0.198)	(0.200)	(0.201)	(0.207)	(0.192)	(0.231)	(0.236)
Course	0.093	0.296	−0.072	−0.045	0.025	0.119	0.043	−0.189	−0.136	−0.240
	(0.205)	(0.213)	(0.205)	(0.202)	(0.203)	(0.205)	(0.211)	(0.196)	(0.235)	(0.240)
Verification	0.114	0.036	0.168	0.178	0.005	0.144	0.145	−0.066	−0.099	−0.034
	(0.217)	(0.225)	(0.216)	(0.213)	(0.215)	(0.217)	(0.223)	(0.207)	(0.249)	(0.254)
Control mean	0.099	4.406	4.514	4.732	4.601	4.623	4.638	0.102	3.913	3.87
Control std. dev.	0.82	0.917	0.938	0.76	0.909	0.873	0.854	0.958	1.229	1.145
Outcome range	[-4,0.8]	[1,5]	[1,5]	[1,5]	[1,5]	[1,5]	[1,5]	[-2.8,1.2]	[1,5]	[1,5]
Observations	463	463	463	463	463	463	463	463	463	463
R^2	0.302	0.302	0.341	0.313	0.363	0.330	0.279	0.366	0.396	0.360

Notes: We report estimates from OLS regression including randomization block fixed effects. Robust standard errors are in parentheses. †: imagine you received an that came to you on WhatsApp with a screenshot of the Twitter account of a well-known person saying something very controversial. ††: imagine you received an audio that came to you by WhatsApp reporting evidence of corruption of a politician that you already suspected and you did not like. * denotes p<0.1, ** denotes p<0.05, and *** denotes p<0.01.

Table S8: Balance on mechanisms: Trust and Attention

	Trust traditional sources index of (1,1,1)	Radio / TV	Newspaper	Internet	Distrust social media sources index of (-1,-1,-1)	WhatsApp	Social media	Conversations	Attention to misinformation index of (1)	How often do you question wether a news is false
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A.1 Sample course Delivery Online										
Course and verification	−0.496***	−0.255	−0.610***	−0.438**	0.186	0.044	−0.149	−0.298	0.120	0.098
	(0.173)	(0.207)	(0.187)	(0.206)	(0.190)	(0.204)	(0.214)	(0.214)	(0.205)	(0.166)
Course	−0.387**	−0.308	−0.450*	−0.276	0.455**	−0.247	−0.495**	−0.359	0.148	0.120
	(0.196)	(0.245)	(0.239)	(0.216)	(0.205)	(0.214)	(0.246)	(0.229)	(0.188)	(0.152)
Verification	0.082	0.124	0.144	−0.013	0.278*	−0.180	−0.200	−0.297	0.084	0.068
	(0.148)	(0.183)	(0.147)	(0.197)	(0.165)	(0.181)	(0.192)	(0.198)	(0.182)	(0.147)
A.2 Sample Course Delivery WhatsApp										
Course and verification	−0.351*	−0.325	−0.417*	−0.212	−0.023	0.149	0.094	−0.141	0.251	0.203
	(0.201)	(0.237)	(0.219)	(0.227)	(0.195)	(0.209)	(0.222)	(0.210)	(0.185)	(0.150)
Course	−0.190	−0.162	−0.246	−0.114	−0.183	0.109	0.164	0.172	0.074	0.060
	(0.205)	(0.242)	(0.224)	(0.232)	(0.199)	(0.213)	(0.227)	(0.214)	(0.189)	(0.153)
Verification	−0.047	0.102	−0.111	−0.104	−0.089	−0.065	0.248	0.003	0.121	0.098
	(0.217)	(0.256)	(0.237)	(0.245)	(0.210)	(0.226)	(0.240)	(0.226)	(0.200)	(0.162)
Control mean	0.12	3.645	3.957	2.746	-0.059	2.072	2.5	2.942	-0.11	3.812
Control std. dev.	0.931	1.113	1.052	1.184	1.016	1.118	1.116	1.151	1.081	0.876
Outcome range	[-2.6,2.1]	[1,5]	[1,5]	[1,5]	[-3,1.8]	[1,5]	[1,5]	[1,5]	[-3.6,1.4]	[1,5]
Observations	463	463	463	463	463	463	463	463	463	463
R^2	0.302	0.274	0.329	0.349	0.345	0.368	0.292	0.357	0.409	0.409

Notes: We report estimates from OLS regression including randomization block fixed effects. Robust standard errors are in parentheses. * denotes p<0.1, ** denotes p<0.05, and *** denotes p<0.01.

Table S9: Balance on knowledge to identify information

	Increase knowledge to identify information index of (1,1,1,1, 1,1,-1,-1,-1)	How much knowledge do you have to identify whether a news is false or not?	Knowing how to identify a fake new index	Radio / TV	Newspaper	Internet	WhatsApp	Social media	Conversations
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A.1 Sample course Delivery Online									
Course and verification	−0.322*	−0.138	0.067	0.104	0.030	0.179**	−0.014	−0.083	0.044
	(0.175)	(0.138)	(0.052)	(0.065)	(0.065)	(0.091)	(0.065)	(0.063)	(0.091)
Course	−0.529**	−0.074	0.001	0.146*	0.215**	0.128	−0.177**	−0.097	−0.00002
	(0.214)	(0.160)	(0.057)	(0.079)	(0.085)	(0.099)	(0.090)	(0.069)	(0.101)
Verification	−0.146	0.015	−0.060	−0.012	0.005	−0.006	−0.139*	−0.003	−0.086
	(0.168)	(0.127)	(0.049)	(0.054)	(0.052)	(0.085)	(0.071)	(0.052)	(0.084)
A.2 Sample Course Delivery WhatsApp									
Course and verification	0.026	0.103	0.021	0.028	0.006	0.038	0.012	0.035	−0.019
	(0.185)	(0.144)	(0.056)	(0.068)	(0.067)	(0.095)	(0.070)	(0.055)	(0.090)
Course	−0.030	0.025	−0.064	−0.062	−0.130*	0.045	−0.021	−0.011	−0.035
	(0.189)	(0.146)	(0.057)	(0.069)	(0.069)	(0.097)	(0.071)	(0.056)	(0.092)
Verification	0.390*	0.242	−0.006	−0.087	−0.066	−0.060	0.100	0.089	0.037
	(0.200)	(0.155)	(0.060)	(0.073)	(0.072)	(0.102)	(0.075)	(0.059)	(0.097)
Control mean	0.035	3.116	0.374	0.159	0.152	0.609	0.862	0.899	0.63
Control std. dev.	1.02	0.855	0.298	0.367	0.36	0.49	0.346	0.303	0.484
Outcome range	[-3.3,2.4]	[1,5]	[-0.4,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]
Observations	463	463	463	463	463	463	463	463	463
R^2	0.408	0.516	0.414	0.366	0.364	0.308	0.373	0.369	0.412

Notes: We report estimates from OLS regression including randomization block fixed effects. Robust standard errors are in parentheses. * denotes p<0.1, ** denotes p<0.05, and *** denotes p<0.01.

Table S10: Balance on knowledge to verify information and consumption behavior

	Increase knowledge to verify information index of (1,1,1)	How much knowledge do you have to verify if a doubtful news is false or not?	Main ways to verify news index	How many verifiers or fact-checkers do you know?	Increase consumption behavior of traditional sources index of (1,1,1)	Radio / TV	Newspaper	Internet	Decrease consumption behavior of social media sources index of (-1,-1,-1)	WhatsApp	Social media
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
A.1 Sample course Delivery Online											
Course and verification	−0.048	−0.068	−0.052	0.179	0.135	−0.123	−0.086	−0.195	−0.449	−0.055	−0.511*
	(0.172)	(0.122)	(0.070)	(0.214)	(0.199)	(0.240)	(0.245)	(0.299)	(0.295)	(0.171)	(0.275)
Course	−0.341*	−0.136	−0.119*	−0.182	0.121	0.197	−0.035	−0.342	−0.093	−0.212	−0.363
	(0.187)	(0.140)	(0.072)	(0.230)	(0.238)	(0.260)	(0.293)	(0.264)	(0.328)	(0.203)	(0.336)
Verification	−0.172	−0.084	−0.071	−0.029	−0.131	−0.396	0.097	0.006	0.486*	−0.121	−0.009
	(0.159)	(0.123)	(0.058)	(0.186)	(0.187)	(0.251)	(0.233)	(0.268)	(0.290)	(0.157)	(0.215)
A.2 Sample Course Delivery WhatsApp											
Course and verification	0.140	0.121	0.064	−0.074	−0.063	−0.038	−0.631**	−0.187	−0.268	−0.222	−0.149
	(0.170)	(0.137)	(0.068)	(0.203)	(0.194)	(0.259)	(0.271)	(0.273)	(0.311)	(0.175)	(0.248)
Course	0.017	−0.024	0.002	0.065	−0.268	−0.006	−0.745***	−0.136	−0.100	−0.148	0.315
	(0.173)	(0.140)	(0.069)	(0.207)	(0.198)	(0.264)	(0.277)	(0.279)	(0.317)	(0.179)	(0.253)
Verification	0.075	0.150	0.022	−0.111	−0.338	−0.357	−0.536*	0.095	0.118	0.095	0.175
	(0.183)	(0.148)	(0.073)	(0.219)	(0.209)	(0.279)	(0.292)	(0.294)	(0.335)	(0.189)	(0.268)
Control mean	0.002	3.109	0.57	1.21	0.008	4.428	4.536	4.094	4.188	5.355	4.355
Control std. dev.	0.993	0.799	0.349	1.13	0.988	1.464	1.446	1.474	1.774	0.965	1.328
Outcome range	[-3.1,2.5]	[1,5]	[-0.3,1]	[0,3]	[-3.1,3.6]	[1,6]	[1,6]	[1,6]	[1,6]	[1,6]	[1,6]
Observations	463	463	463	463	463	463	463	463	463	463	463
R^2	0.503	0.530	0.306	0.454	0.349	0.429	0.404	0.424	0.457	0.494	0.412

Notes: We report estimates from OLS regression including randomization block fixed effects. Robust standard errors are in parentheses. * denotes p<0.1, ** denotes p<0.05, and *** denotes p<0.01.

Table S11: Balance on sharing and verifying behavior

	Decrease sharing behavior index of (1,-1, -1,-1)	If you know a story is false, how often do you share its falsehood with others on social media?	Share scenario 1 withoout verification†	Share scenario 3 withoout verification††	Increase verifying behavior index of (1)	How often do you verify news that you doubt may be false before sharing it?
	(1)	(2)	(3)	(4)	(5)	(6)
A.1 Sample course Delivery Online						
Course and verification	−0.049	−0.012	0.148	−0.065	0.030	0.028
	(0.181)	(0.215)	(0.196)	(0.209)	(0.194)	(0.180)
Course	−0.181	−0.342	−0.076	0.046	0.070	0.065
	(0.201)	(0.235)	(0.204)	(0.282)	(0.205)	(0.190)
Verification	0.079	0.361*	0.105	0.208	−0.128	−0.119
	(0.173)	(0.203)	(0.177)	(0.217)	(0.196)	(0.183)
A.2 Sample Course Delivery WhatsApp						
Course and verification	−0.090	−0.231	0.075	−0.198	0.088	0.082
	(0.177)	(0.228)	(0.194)	(0.229)	(0.194)	(0.181)
Course	−0.190	−0.029	0.332*	0.058	0.000	0.000
	(0.180)	(0.232)	(0.197)	(0.233)	(0.198)	(0.184)
Verification	0.016	0.103	0.241	−0.166	0.192	0.179
	(0.191)	(0.246)	(0.209)	(0.247)	(0.209)	(0.195)
Control mean	-0.02	2.768	1.826	2.152	-0.011	3.797
Control std. dev.	1.017	1.192	1.039	1.317	1.072	0.998
Outcome range	[-3,2.1]	[1,5]	[1,5]	[1,5]	[-3,1.3]	[1,5]
Observations	463	463	463	463	463	463
R^2	0.461	0.394	0.436	0.392	0.350	0.350

Notes: We report estimates from OLS regression including randomization block fixed effects. Robust standard errors are in parentheses. †: imagine you received an that came to you on WhatsApp with a screenshot of the Twitter account of a well-known person saying something very controversial. ††: imagine you received an audio that came to you by WhatsApp reporting evidence of corruption of a politician that you already suspected and you did not like. * denotes p<0.1, ** denotes p<0.05, and *** denotes p<0.01.

Treatment Effects

Table S12: Endline survey questions used to create all outcome indices

Indexes	Variables	Direction
Misinformation importance index	How likely do you think misinformation contributes to the following problems in Bolivian society?[1]	1
Likelihood of false of traditional sources index	How much do you think the information in these sources is true? - Traditional media[2]	-1
Likelihood of false of social media sources index	How much do you think the information in these sources is true? - Social media[2]	1
	Hypothetical scenario 1: How likely is it that you think it's fake?[3]	1
	Hypothetical scenario 3: How likely is it that you think it's fake?[3]	1
Distrust traditional sources index	How much do you trust the information you receive from these sources? - Traditional media[2]	1
Distrust social media sources index	How much do you trust the information you receive from these sources? - Social media[2]	-1
Attention to misinformation index	How often do you question whether a news is false	1
Knowledge to identify information index	How much knowledge do you have to identify whether a news is false or not?	1
	What are the main characteristics of a piece of news that make you doubt that it may be false?	1
	How many recent misinformation cases in Bolivia do you know?	1
	What are the main means by which fake news spread? - Traditional media[2]	-1
	What are the main means by which fake news spread? - Social media[2]	1
Knowledge to verify information index	How much knowledge do you have to verify if doubtful news is false or not?	1
	What are the main ways to verify news?	1
	If you wanted to verify some information, what fact-checkers could you use? - Number of fact-checkers	1
Consumption behavior of traditional sources index	How often do you get news from these different types of sources? - Traditional media[2]	1
Consumption behavior of social media sources index	How often do you get news from these different types of sources? - Social media[2]	-1
Sharing behavior index	If you know a story is false, how often do you share its falsehood with others on social media?	1
	How often do you share news you receive on WhatsApp and social media	-1
	Hypothetical scenario 1: How likely are you to share it without verifying it first?[3]	-1
	Hypothetical scenario 3: How likely are you to share it without verifying it first?[3]	-1
Verifying behavior index	How often do you verify news that you doubt may be false before sharing it?	1

Notes: [1] Problems in Bolivian society: Decisions that may affect health, the election of candidates who do not represent the interests of citizens, disparage or exalt people, ideological polarization, violence towards certain people or falsely from society, and increase hatred towards certain people or groups in society.

[2] Traditional media: Radio/TV, Newspaper, and webpages. Social media: WhatsApp, Conversations with family and friends, WhatsApp with family and friends, and WhatsApp group with unknown people. [3] Hypothetical scenario 1: You received an image that came to you on WhatsApp with a screenshot of the Twitter account of a well-known person saying something very controversial.

Hypothetical scenario 3: You received an audio that came to you by WhatsApp reporting evidence of corruption of a politician that you already suspected and you did not like.

Table S13: Treatment effects on indexes by course delivered online and via WhatsApp

Panel A: Treatment effects with covariates in the outcome family

	Mis-information importance index (1)	Decrease likelihood of false of traditional sources index (2)	Increase likelihood of false of social media sources index (3)	Trust traditional sources index (4)	Disrust social media sources index (5)	Attention to mis-information index (6)	Increase knowledge to identify information index (7)	Increase knowledge to verify information index (8)	Increase consumption behavior of traditional sources index (9)	Decrease consumption behavior of social media sources index (10)	Decrease sharing behavior index (11)	Increase verifying behavior index (12)
A.1 Sample course Delivery Online												
Course and verification	−0.004	0.290	−0.116	−0.036	−0.046	0.218	0.197	0.606***	0.054	−0.228	0.152	0.135
	(0.188)	(0.180)	(0.153)	(0.178)	(0.187)	(0.173)	(0.174)	(0.133)	(0.154)	(0.165)	(0.184)	(0.175)
Course	0.148	−0.084	0.173	−0.203	0.075	0.090	0.245	0.569***	−0.148	−0.088	0.292	0.190
	(0.209)	(0.196)	(0.169)	(0.177)	(0.188)	(0.171)	(0.193)	(0.160)	(0.177)	(0.157)	(0.232)	(0.197)
Verification	−0.039	−0.009	−0.193	−0.126	−0.083	0.226	−0.036	−0.004	0.143	−0.258*	0.374**	−0.012
	(0.167)	(0.154)	(0.162)	(0.164)	(0.190)	(0.155)	(0.160)	(0.117)	(0.130)	(0.151)	(0.184)	(0.179)
A.2 Sample Course Delivery WhatsApp												
Course and verification	−0.023	0.073	0.357**	0.097	0.109	0.297**	0.076	0.154	−0.191	0.269*	−0.023	0.038
	(0.145)	(0.144)	(0.140)	(0.138)	(0.137)	(0.127)	(0.145)	(0.128)	(0.159)	(0.145)	(0.143)	(0.144)
Course	−0.277*	−0.184	0.262*	0.050	0.343**	0.162	0.072	0.215*	−0.081	0.364***	0.102	0.049
	(0.166)	(0.144)	(0.152)	(0.140)	(0.149)	(0.134)	(0.144)	(0.120)	(0.171)	(0.120)	(0.144)	(0.152)
Verification	−0.184	0.053	0.285**	0.002	0.227*	0.114	0.025	0.087	−0.069	0.464***	0.243	0.208
	(0.168)	(0.148)	(0.131)	(0.151)	(0.136)	(0.129)	(0.165)	(0.133)	(0.154)	(0.138)	(0.169)	(0.131)
R^2	0.379	0.462	0.485	0.491	0.456	0.543	0.450	0.604	0.457	0.531	0.381	0.437

Panel B: IPSW with covariates in the outcome family

Panel B.1: Sample course Delivery Online												
Course and verification	0.002	0.275	−0.115	−0.026	−0.038	0.247	0.186	0.624***	0.059	−0.218	0.142	0.151
	(0.186)	(0.179)	(0.155)	(0.178)	(0.182)	(0.175)	(0.172)	(0.138)	(0.160)	(0.170)	(0.179)	(0.174)
Course	0.174	−0.036	0.175	−0.185	0.093	0.118	0.228	0.643***	−0.199	−0.055	0.366*	0.259
	(0.211)	(0.197)	(0.175)	(0.175)	(0.181)	(0.168)	(0.190)	(0.160)	(0.182)	(0.158)	(0.198)	(0.189)
Verification	−0.076	−0.029	−0.149	−0.148	−0.048	0.271*	−0.008	0.028	0.146	−0.264*	0.361**	−0.007
	(0.170)	(0.161)	(0.170)	(0.171)	(0.189)	(0.161)	(0.165)	(0.124)	(0.137)	(0.155)	(0.180)	(0.181)
B.2 Sample Course Delivery WhatsApp												
Course and verification	−0.084	0.096	0.332**	0.096	0.112	0.322**	0.068	0.109	−0.169	0.303**	−0.018	0.025
	(0.162)	(0.150)	(0.143)	(0.140)	(0.140)	(0.131)	(0.147)	(0.128)	(0.162)	(0.141)	(0.143)	(0.147)
Course	−0.343**	−0.185	0.245	0.063	0.358**	0.178	0.065	0.221*	−0.041	0.388***	0.082	0.038
	(0.174)	(0.155)	(0.148)	(0.143)	(0.149)	(0.133)	(0.147)	(0.120)	(0.171)	(0.120)	(0.143)	(0.158)
Verification	−0.163	0.117	0.275**	0.042	0.199	0.133	−0.015	0.064	−0.033	0.489***	0.278*	0.189
	(0.176)	(0.146)	(0.132)	(0.149)	(0.136)	(0.131)	(0.158)	(0.128)	(0.165)	(0.140)	(0.163)	(0.144)
R^2	0.385	0.499	0.519	0.533	0.485	0.589	0.487	0.620	0.484	0.561	0.437	0.499

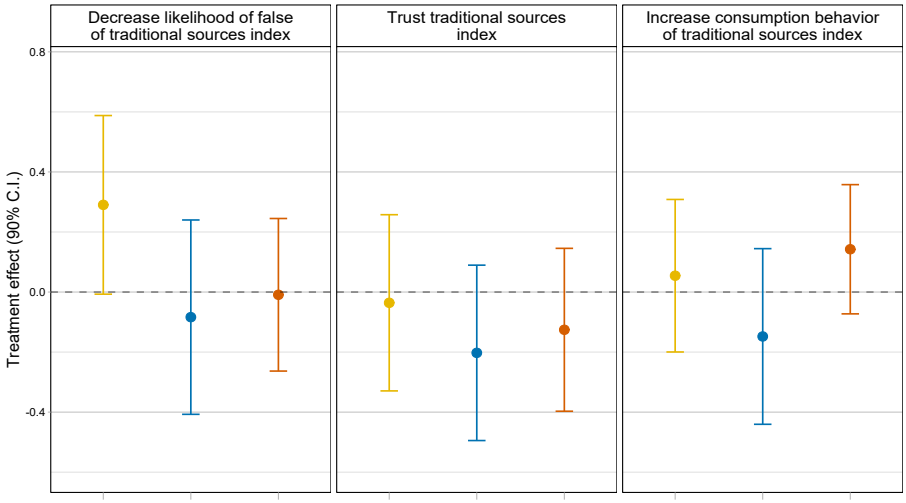
Panel C: Treatment effects with covariates selected with LASSO

Panel C.1: Sample course Delivery Online												
Course and verification	−0.029	0.393**	−0.189	0.028	0.003	0.195	0.212	0.573***	0.115	−0.265	0.118	0.064
	(0.189)	(0.169)	(0.156)	(0.179)	(0.183)	(0.172)	(0.161)	(0.135)	(0.152)	(0.167)	(0.172)	(0.157)
Course	0.147	−0.042	0.221	−0.180	0.097	0.116	0.374**	0.634***	−0.116	−0.072	0.331	0.180
	(0.202)	(0.193)	(0.174)	(0.175)	(0.187)	(0.172)	(0.183)	(0.140)	(0.164)	(0.161)	(0.215)	(0.181)
Verification	−0.041	−0.043	−0.194	−0.113	−0.128	0.220	0.039	−0.002	0.162	−0.264*	0.423**	−0.069
	(0.167)	(0.147)	(0.159)	(0.157)	(0.186)	(0.154)	(0.148)	(0.111)	(0.124)	(0.147)	(0.169)	(0.164)
C.2 Sample Course Delivery WhatsApp												
Course and verification	−0.059	0.064	0.327**	0.149	0.165	0.267**	0.092	0.193	−0.107	0.255*	−0.022	0.045
	(0.145)	(0.130)	(0.131)	(0.135)	(0.136)	(0.129)	(0.137)	(0.120)	(0.150)	(0.144)	(0.138)	(0.139)
Course	−0.301*	−0.207	0.234	0.102	0.390***	0.119	0.129	0.235**	−0.082	0.404***	0.131	0.100
	(0.164)	(0.134)	(0.146)	(0.134)	(0.142)	(0.143)	(0.128)	(0.116)	(0.161)	(0.117)	(0.146)	(0.142)
Verification	−0.192	0.047	0.279**	0.013	0.266*	0.099	0.080	0.045	−0.030	0.443***	0.319*	0.194
	(0.168)	(0.136)	(0.126)	(0.146)	(0.136)	(0.133)	(0.152)	(0.127)	(0.151)	(0.137)	(0.170)	(0.128)
Outcome mean	0	0	0	0	0	0	0	0	0	0	0	0
Outcome std. dev.	1	1	1	1	1	1	1	1	1	1	1	1
Outcome range	[-3.2,0.6]	[-4.1,3.1]	[-3.8,2.5]	[-2.6,2.2]	[-3.4,1.7]	[-3.6,1.4]	[-3.3,2.5]	[-3.3,2.1]	[-2.8,1.8]	[-1.7,3.3]	[-3,5.1]	[-3,1.3]
Num. LASSO covariates	0	9	8	2	1	0	8	7	12	3	6	7
Observations	463	463	463	463	463	463	463	463	463	463	463	463
R^2	0.367	0.531	0.500	0.508	0.451	0.523	0.537	0.656	0.515	0.538	0.441	0.506

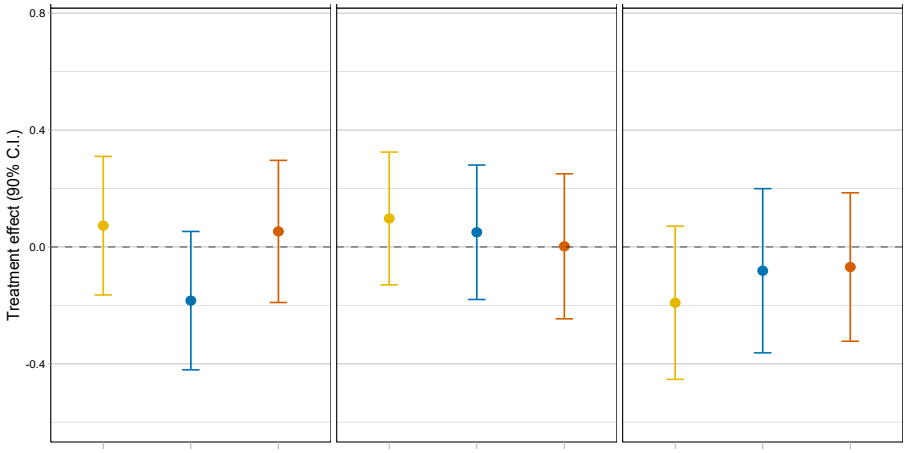
Notes: We report estimates from OLS regression including randomization block fixed effects. All specifications in all panels include the corresponding outcome variables at baseline as a control. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel C are from a 10-fold cross validation LASSO model with lambda chosen to be that of the minimum average cross-validated error. The treatment indicators, lagged dependent variable, and fixed effects are forced into the LASSO model and covariates are selected from a pool of 72 baseline variables for each specification. Robust standard errors are in parentheses. * denotes $p<0.1$, ** denotes $p<0.05$, and *** denotes $p<0.01$.

Figure S1: Treatment effects on traditional sources: likelihood of false, distrust, and consumption of traditional sources

(a) Sample course delivery online



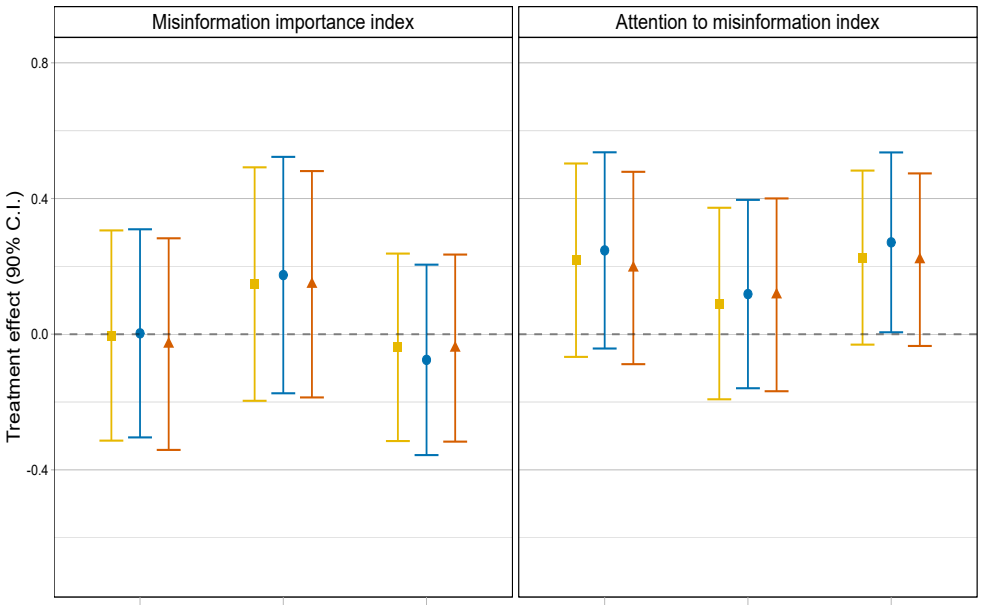
(b) Sample course delivery WhatsApp



Notes: The estimates and 90% confidence intervals in each box are from separate OLS regressions. The labels are the corresponding dependent variables regressed on treatment indicators and randomization block fixed effects. The outcomes are in columns (2), (4), and (9) of Panel A of Appendix Table S13.

Figure S2: Treatment effects comparison between OLS, IPSW, and LASSO methods on mechanisms: misinformation importance and attention to misinformation

(a) Sample course delivery online



(b) Sample course delivery WhatsApp

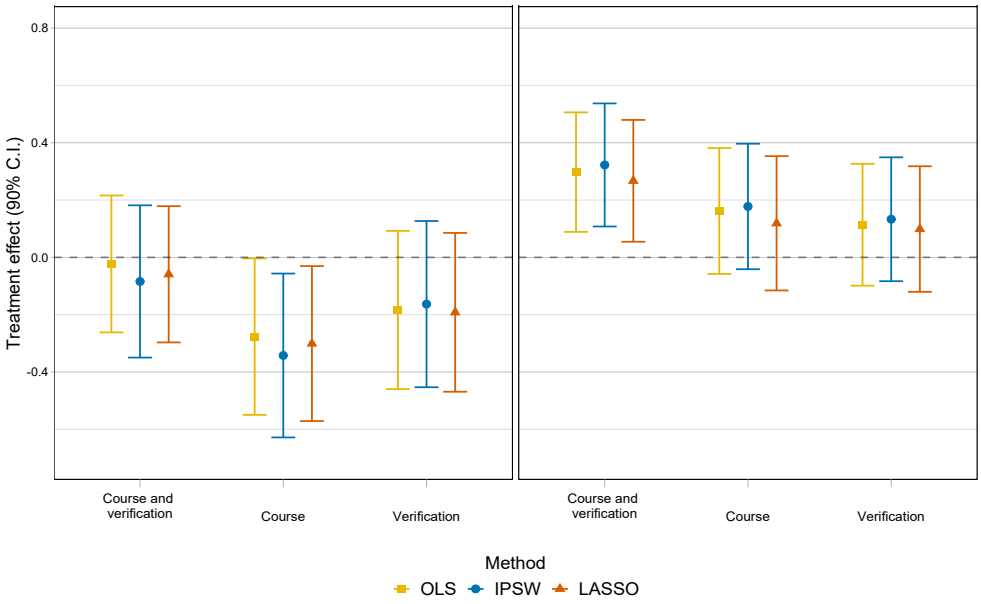
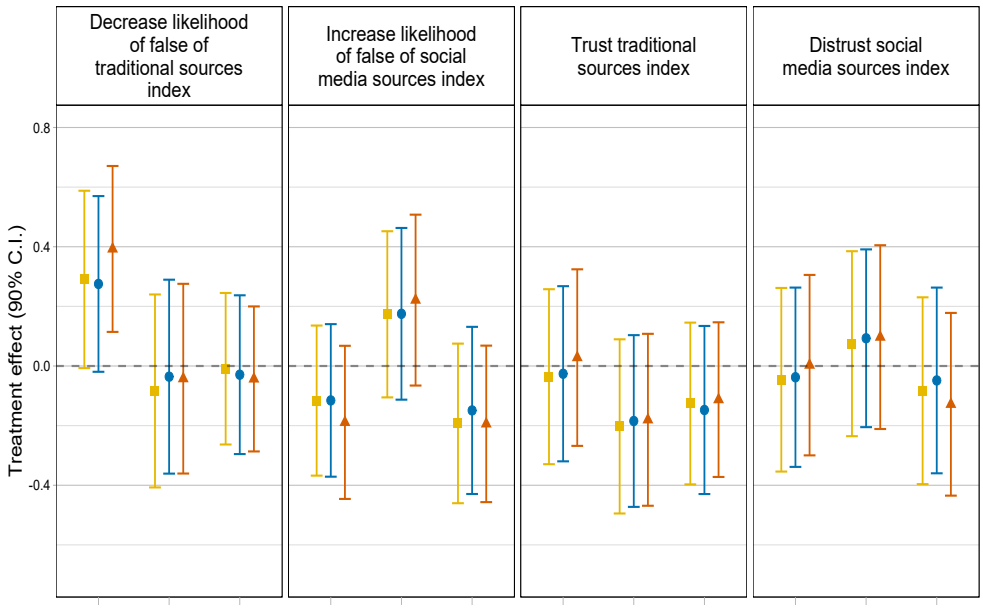


Figure S3: Treatment effects comparison between OLS, IPSW, and LASSO methods on mechanisms: likelihood of false and distrust of traditional and social media sources

(a) Sample course delivery online



(b) Sample course delivery WhatsApp

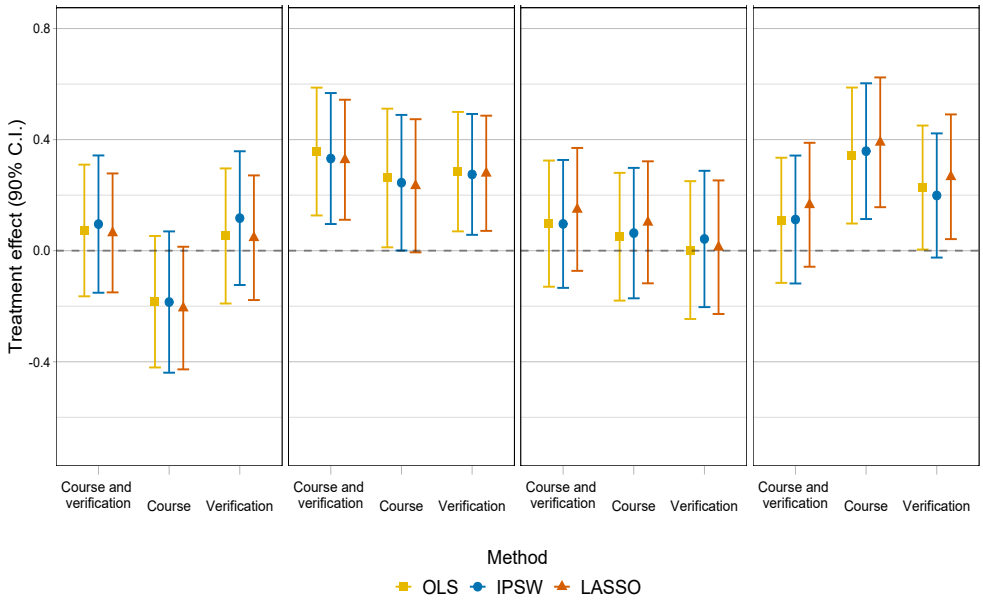
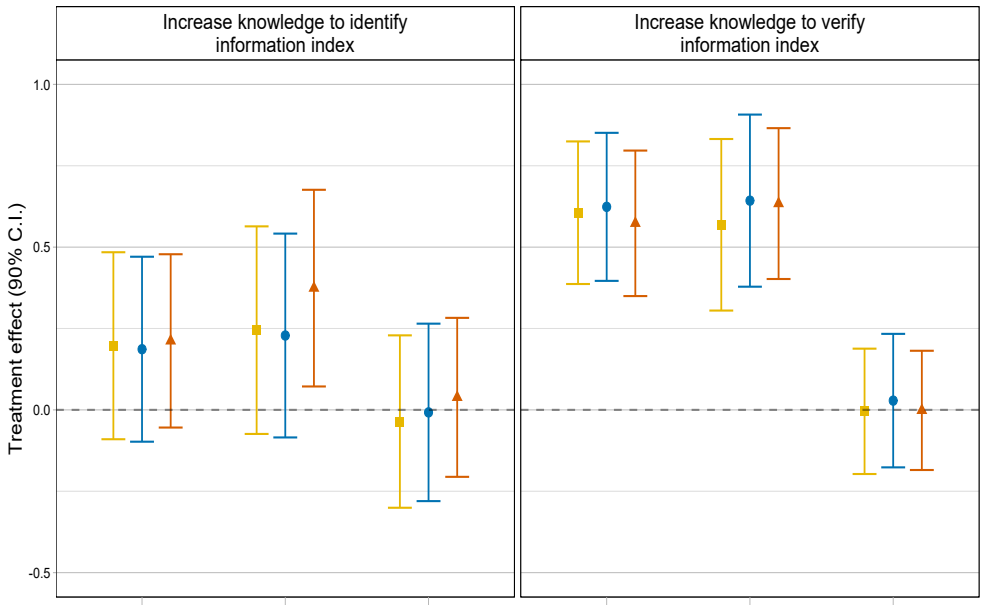


Figure S4: Treatment effects comparison between OLS, IPSW, and LASSO methods on knowledge to identify and verify information

(a) Sample course delivery online



(b) Sample course delivery WhatsApp

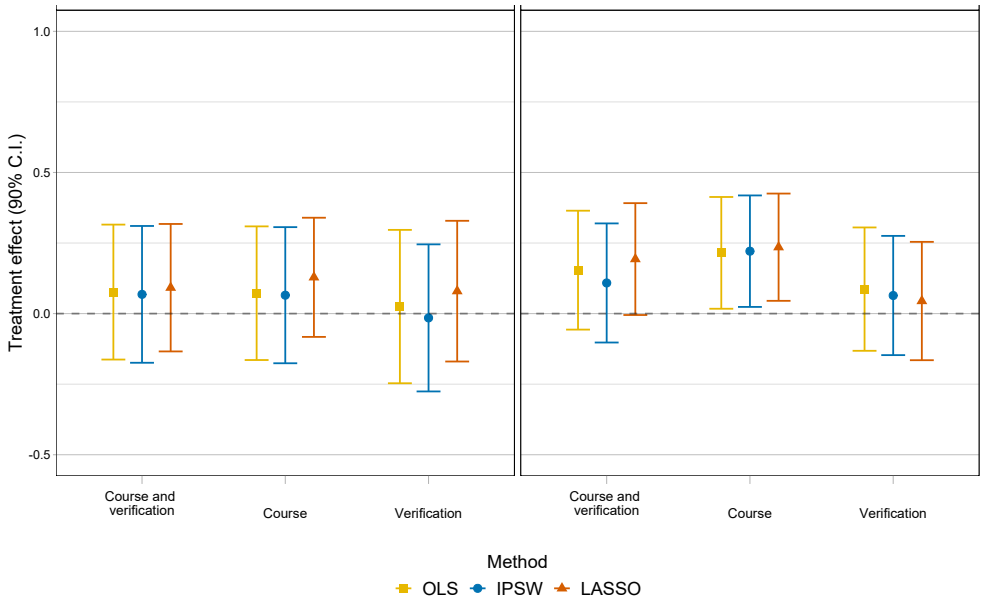
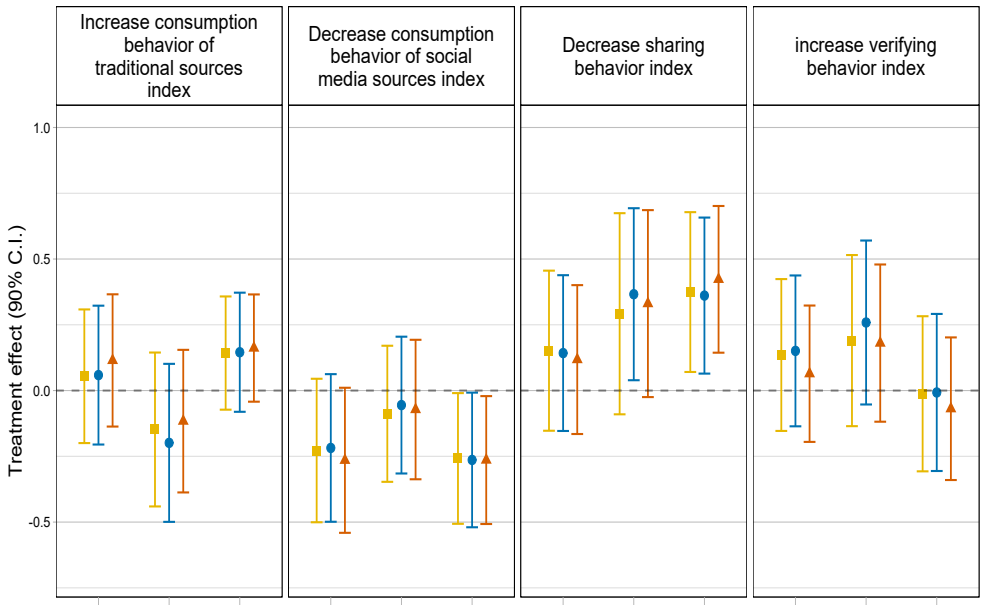


Figure S5: Treatment effects comparison between OLS, IPSW, and LASSO methods on consumption, sharing, and verifying behavior

(a) Sample course delivery online



(b) Sample course delivery WhatsApp

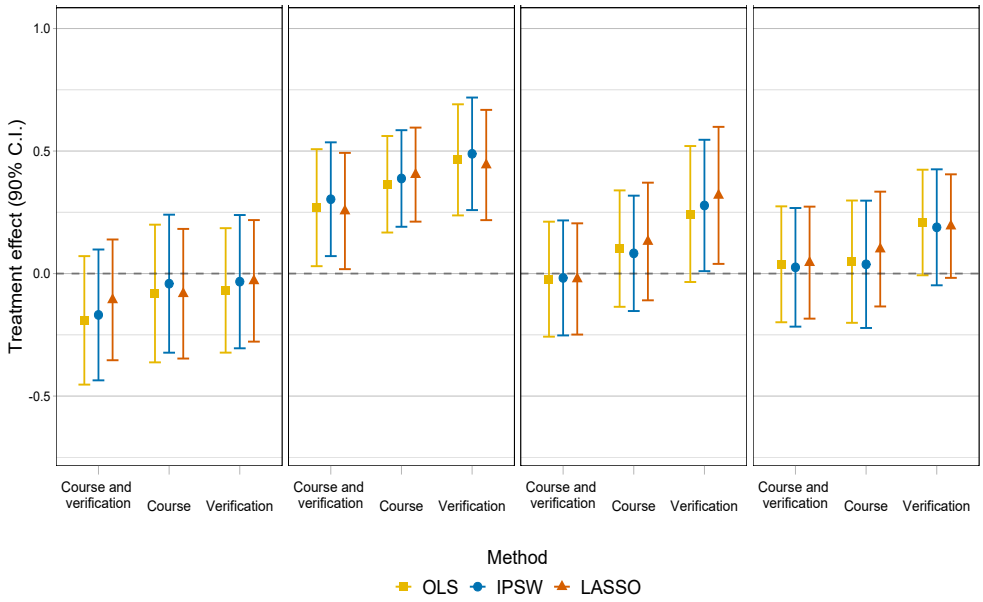


Table S14: Treatment effects on mechanisms: Misinformation relevance index's components

	Misinformation importance index of (1,1,1,1 1,1,1) (1)	Decisions that may affect health (2)	Election of candidates who do not represent the interests of citizens (3)	To falsely discredit or glorify people (4)	Ideological polarization (5)	Violence towards certain people or groups in society (6)	Increase hatred towards certain people or groups in society (7)
Panel A: Treatment effects with covariates in the outcome family							
A.1 Sample course Delivery Online							
Course and verification	−0.004	0.056	−0.165	−0.003	0.109	0.168	−0.041
	(0.188)	(0.222)	(0.235)	(0.226)	(0.217)	(0.212)	(0.219)
Course	0.148	0.075	0.239	0.062	0.116	0.114	0.192
	(0.209)	(0.259)	(0.248)	(0.227)	(0.232)	(0.236)	(0.255)
Verification	−0.039	0.044	−0.100	−0.027	−0.058	−0.011	−0.193
	(0.167)	(0.215)	(0.199)	(0.187)	(0.189)	(0.202)	(0.207)
A.2 Sample Course Delivery WhatsApp							
Course and verification	−0.023	−0.045	0.092	0.070	0.114	−0.081	−0.185
	(0.145)	(0.162)	(0.181)	(0.171)	(0.178)	(0.164)	(0.170)
Course	−0.277*	−0.381**	−0.174	−0.126	−0.147	−0.440**	−0.293
	(0.166)	(0.184)	(0.199)	(0.186)	(0.194)	(0.192)	(0.184)
Verification	−0.184	−0.525***	−0.059	0.061	0.120	−0.128	−0.338*
	(0.168)	(0.199)	(0.207)	(0.185)	(0.207)	(0.182)	(0.190)
R^2	0.379	0.339	0.383	0.404	0.394	0.410	0.406
Panel B: IPSW with covariates in the outcome family							
Panel B.1: Sample course Delivery Online							
Course and verification	0.002	0.058	−0.194	0.029	0.168	0.184	−0.054
	(0.186)	(0.227)	(0.235)	(0.221)	(0.215)	(0.210)	(0.221)
Course	0.174	0.108	0.237	0.103	0.126	0.161	0.238
	(0.211)	(0.268)	(0.239)	(0.229)	(0.232)	(0.239)	(0.262)
Verification	−0.076	0.005	−0.149	−0.090	−0.065	−0.055	−0.259
	(0.170)	(0.232)	(0.200)	(0.191)	(0.193)	(0.203)	(0.211)
B.2 Sample Course Delivery WhatsApp							
Course and verification	−0.084	−0.107	0.014	0.003	0.040	−0.136	−0.243
	(0.162)	(0.177)	(0.196)	(0.186)	(0.197)	(0.179)	(0.183)
Course	−0.343**	−0.451**	−0.258	−0.214	−0.231	−0.492**	−0.376*
	(0.174)	(0.193)	(0.210)	(0.195)	(0.207)	(0.196)	(0.192)
Verification	−0.163	−0.488**	−0.030	0.094	0.114	−0.080	−0.309
	(0.176)	(0.205)	(0.217)	(0.192)	(0.219)	(0.191)	(0.191)
R^2	0.385	0.355	0.394	0.415	0.386	0.429	0.435
Panel C: Treatment effects with covariates selected with LASSO							
Panel C.1: Sample course Delivery Online							
Course and verification	−0.029	0.058	−0.172	0.081	0.026	0.108	0.018
	(0.189)	(0.218)	(0.233)	(0.220)	(0.230)	(0.208)	(0.217)
Course	0.147	0.031	0.201	−0.004	0.152	0.099	0.135
	(0.202)	(0.255)	(0.232)	(0.200)	(0.235)	(0.232)	(0.227)
Verification	−0.041	0.062	−0.144	−0.077	−0.099	−0.005	−0.191
	(0.167)	(0.207)	(0.185)	(0.169)	(0.193)	(0.203)	(0.195)
C.2 Sample Course Delivery WhatsApp							
Course and verification	−0.059	−0.076	0.068	0.119	0.055	−0.160	−0.184
	(0.145)	(0.160)	(0.175)	(0.173)	(0.175)	(0.166)	(0.171)
Course	−0.301*	−0.348*	−0.228	−0.078	−0.217	−0.531***	−0.181
	(0.164)	(0.181)	(0.196)	(0.186)	(0.191)	(0.197)	(0.185)
Verification	−0.192	−0.482**	−0.092	0.062	0.098	−0.124	−0.321*
	(0.168)	(0.200)	(0.201)	(0.184)	(0.210)	(0.183)	(0.191)
Outcome mean	0	4.287	4.272	4.46	4.393	4.393	4.43
Outcome std. dev.	1	1.152	1.229	1.163	1.197	1.17	1.22
Outcome range	[-3.2,0.6]	[1,5]	[1,5]	[1,5]	[1,5]	[1,5]	[1,5]
Num. LASSO covariates	0	1	2	11	1	1	6
Observations	463	463	463	463	463	463	463
R^2	0.367	0.345	0.488	0.469	0.372	0.395	0.438

Notes: We report estimates from OLS regression including randomization block fixed effects. All specifications in all panels include the corresponding outcome variables at baseline as a control. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel C are from a 10-fold cross validation LASSO model with lambda chosen to be that of the minimum average cross-validated error. The treatment indicators, lagged dependent variable, and fixed effects are forced into the LASSO model and covariates are selected from a pool of 72 baseline variables for each specification. Robust standard errors are in parentheses. * denotes p<0.1, ** denotes p<0.05, and *** denotes p<0.01.

Table S15: Treatment effects on mechanisms: Likelihood of false index's components

Panel A: Treatment effects with covariates in the outcome family

	Decrease likelihood of false of traditional sources index of (-1,-1,-1)	Radio / TV	Newspaper	Internet	Increase likelihood of false of social media sources index of (1,1,1,1,1)	WhatsApp	Social media	Conversations	Believe scenario 1 is fake†	Believe scenario 3 is fake††
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A.1 Sample course Delivery Online										
Course and verification	0.290	−0.191	−0.179	−0.132	−0.116	0.065	0.120	−0.141	−0.128	−0.249
	(0.180)	(0.151)	(0.142)	(0.156)	(0.153)	(0.128)	(0.131)	(0.126)	(0.254)	(0.220)
Course	−0.084	−0.007	−0.134	0.206	0.173	0.217	0.044	−0.177	0.530**	0.077
	(0.196)	(0.166)	(0.127)	(0.190)	(0.169)	(0.144)	(0.128)	(0.135)	(0.236)	(0.234)
Verification	−0.009	−0.177	−0.073	0.156	−0.193	−0.059	−0.120	−0.338***	0.292	−0.141
	(0.154)	(0.124)	(0.109)	(0.132)	(0.162)	(0.110)	(0.108)	(0.122)	(0.236)	(0.204)
A.2 Sample Course Delivery WhatsApp										
Course and verification	0.073	−0.211**	−0.033	0.042	0.357**	0.193	0.158	0.204*	0.382**	−0.009
	(0.144)	(0.103)	(0.114)	(0.110)	(0.140)	(0.121)	(0.101)	(0.120)	(0.186)	(0.154)
Course	−0.184	0.008	0.205*	0.074	0.262*	0.258**	0.047	0.119	0.276	0.116
	(0.144)	(0.108)	(0.123)	(0.114)	(0.152)	(0.108)	(0.111)	(0.121)	(0.182)	(0.160)
Verification	0.053	−0.127	−0.032	0.023	0.285**	0.222**	0.003	0.082	0.479**	0.123
	(0.148)	(0.119)	(0.124)	(0.107)	(0.131)	(0.105)	(0.117)	(0.111)	(0.204)	(0.180)
R^2	0.462	0.439	0.393	0.380	0.485	0.469	0.420	0.414	0.408	0.452

Panel B: IPSW with covariates in the outcome family

Panel B.1: Sample course Delivery Online										
Course and verification	0.275	−0.220	−0.205	−0.079	−0.115	0.083	0.110	−0.157	−0.080	−0.269
	(0.179)	(0.152)	(0.142)	(0.163)	(0.155)	(0.132)	(0.134)	(0.131)	(0.259)	(0.218)
Course	−0.036	−0.026	−0.186	0.194	0.175	0.192	0.048	−0.171	0.555**	0.047
	(0.197)	(0.162)	(0.125)	(0.195)	(0.175)	(0.145)	(0.134)	(0.140)	(0.238)	(0.229)
Verification	−0.029	−0.172	−0.087	0.188	−0.149	−0.042	−0.103	−0.314**	0.343	−0.137
	(0.161)	(0.132)	(0.112)	(0.144)	(0.170)	(0.120)	(0.114)	(0.129)	(0.242)	(0.211)
B.2 Sample Course Delivery WhatsApp										
Course and verification	0.096	−0.209*	−0.020	0.004	0.332**	0.175	0.137	0.215*	0.339*	−0.013
	(0.150)	(0.107)	(0.119)	(0.112)	(0.143)	(0.124)	(0.103)	(0.122)	(0.194)	(0.162)
Course	−0.185	0.006	0.214*	0.070	0.245	0.244**	0.048	0.131	0.228	0.096
	(0.155)	(0.117)	(0.128)	(0.117)	(0.148)	(0.112)	(0.114)	(0.125)	(0.184)	(0.159)
Verification	0.117	−0.157	−0.055	−0.025	0.275**	0.190*	−0.019	0.101	0.434**	0.167
	(0.146)	(0.128)	(0.126)	(0.103)	(0.132)	(0.111)	(0.115)	(0.114)	(0.208)	(0.180)
R^2	0.499	0.497	0.424	0.426	0.519	0.476	0.452	0.445	0.439	0.501

Panel C: Treatment effects with covariates selected with LASSO

Panel C.1: Sample course Delivery Online										
Course and verification	0.393**	−0.289**	−0.265**	−0.147	−0.189	0.024	0.097	−0.157	−0.024	−0.161
	(0.169)	(0.141)	(0.132)	(0.154)	(0.156)	(0.121)	(0.132)	(0.124)	(0.246)	(0.203)
Course	−0.042	−0.146	−0.246*	0.154	0.221	0.248*	0.009	−0.119	0.642**	0.417*
	(0.193)	(0.172)	(0.132)	(0.183)	(0.174)	(0.137)	(0.126)	(0.133)	(0.250)	(0.222)
Verification	−0.043	−0.196	−0.155	0.154	−0.194	0.020	−0.145	−0.346***	0.383*	0.002
	(0.147)	(0.125)	(0.106)	(0.131)	(0.159)	(0.105)	(0.108)	(0.115)	(0.225)	(0.196)
C.2 Sample Course Delivery WhatsApp										
Course and verification	0.064	−0.250***	−0.007	0.033	0.327**	0.141	0.165	0.120	0.222	0.092
	(0.130)	(0.091)	(0.101)	(0.110)	(0.131)	(0.099)	(0.101)	(0.109)	(0.184)	(0.160)
Course	−0.207	0.041	0.218*	0.049	0.234	0.277***	0.059	0.119	0.144	0.228
	(0.134)	(0.098)	(0.111)	(0.115)	(0.146)	(0.103)	(0.110)	(0.111)	(0.182)	(0.156)
Verification	0.047	−0.083	0.039	−0.000	0.279**	0.227**	0.024	0.019	0.397*	0.192
	(0.136)	(0.107)	(0.113)	(0.108)	(0.126)	(0.102)	(0.114)	(0.105)	(0.204)	(0.180)
Outcome mean	0	2.391	2.244	3.263	0	3.716	3.495	3.168	3.886	3.743
Outcome std. dev.	1	0.806	0.741	0.77	1	0.777	0.744	0.746	1.308	1.212
Outcome range	[-4.1,3.1]	[1,5]	[1,5]	[1,5]	[-3.8,2.5]	[1,5]	[1,5]	[1,5]	[1,5]	[1,5]
Num. LASSO covariates	9	2	11	1	8	15	2	9	55	57
Observations	463	463	463	463	463	463	463	463	463	463
R^2	0.531	0.499	0.495	0.390	0.500	0.555	0.426	0.475	0.527	0.565

Notes: We report estimates from OLS regression including randomization block fixed effects. All specifications in all panels include the corresponding outcome variables at baseline as a control. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel C are from a 10-fold cross validation LASSO model with lambda chosen to be that of the minimum average cross-validated error. The treatment indicators, lagged dependent variable, and fixed effects are forced into the LASSO model and covariates are selected from a pool of 72 baseline variables for each specification. Robust standard errors are in parentheses. †: imagine you received an that came to you on WhatsApp with a screenshot of the Twitter account of a well-known person saying something very controversial. ††: imagine you received an audio that came to you by WhatsApp reporting evidence of corruption of a politician that you already suspected and you did not like. * denotes p<0.1, ** denotes p<0.05, and *** denotes p<0.01.

Table S16: Treatment effects on mechanisms: Distrust index's components

Panel A: Treatment effects with covariates in the outcome family

	Trust traditional sources index of (1,1,1) (1)	Radio / TV (2)	Newspaper (3)	Internet (4)	Distrust social media sources index of (-1,-1,-1) (5)	WhatsApp (6)	Social media (7)	Conversations (8)
A.1 Sample course Delivery Online								
Course and verification	−0.036	−0.161	0.122	−0.010	−0.046	0.058	0.111	−0.039
	(0.178)	(0.209)	(0.212)	(0.184)	(0.187)	(0.184)	(0.200)	(0.165)
Course	−0.203	0.006	0.352*	−0.625***	0.075	0.099	−0.069	−0.186
	(0.177)	(0.217)	(0.205)	(0.206)	(0.188)	(0.190)	(0.207)	(0.174)
Verification	−0.126	0.126	−0.136	−0.216	−0.083	0.266	−0.046	0.014
	(0.164)	(0.180)	(0.175)	(0.184)	(0.190)	(0.187)	(0.189)	(0.175)
A.2 Sample Course Delivery WhatsApp								
Course and verification	0.097	0.294*	−0.103	0.111	0.109	−0.306**	0.108	−0.074
	(0.138)	(0.166)	(0.161)	(0.165)	(0.137)	(0.128)	(0.141)	(0.151)
Course	0.050	0.112	0.012	0.038	0.343**	−0.299**	−0.172	−0.347**
	(0.140)	(0.163)	(0.159)	(0.171)	(0.149)	(0.129)	(0.151)	(0.166)
Verification	0.002	0.155	0.065	−0.106	0.227*	−0.232*	0.032	−0.346**
	(0.151)	(0.170)	(0.145)	(0.188)	(0.136)	(0.127)	(0.157)	(0.152)
R^2	0.491	0.456	0.451	0.464	0.456	0.418	0.449	0.480

Panel B: IPSW with covariates in the outcome family

Panel B.1: Sample course Delivery Online								
Course and verification	−0.026	−0.108	0.183	−0.072	−0.038	0.020	0.123	−0.023
	(0.178)	(0.212)	(0.209)	(0.186)	(0.182)	(0.180)	(0.198)	(0.161)
Course	−0.185	0.071	0.423**	−0.679***	0.093	0.032	−0.033	−0.200
	(0.175)	(0.216)	(0.199)	(0.214)	(0.181)	(0.184)	(0.204)	(0.167)
Verification	−0.148	0.150	−0.126	−0.272	−0.048	0.197	−0.069	0.034
	(0.171)	(0.190)	(0.179)	(0.190)	(0.189)	(0.187)	(0.193)	(0.175)
B.2 Sample Course Delivery WhatsApp								
Course and verification	0.096	0.291*	−0.152	0.140	0.112	−0.322**	0.089	−0.036
	(0.140)	(0.172)	(0.171)	(0.167)	(0.140)	(0.127)	(0.147)	(0.157)
Course	0.063	0.120	−0.014	0.074	0.358**	−0.310**	−0.205	−0.343**
	(0.143)	(0.170)	(0.166)	(0.174)	(0.149)	(0.128)	(0.152)	(0.170)
Verification	0.042	0.185	0.038	−0.030	0.199	−0.258**	0.058	−0.276*
	(0.149)	(0.173)	(0.151)	(0.190)	(0.136)	(0.124)	(0.151)	(0.160)
R^2	0.533	0.485	0.487	0.498	0.485	0.441	0.483	0.512

Panel C: Treatment effects with covariates selected with LASSO

Panel C.1: Sample course Delivery Online								
Course and verification	0.028	−0.088	0.087	0.051	0.003	0.071	0.050	−0.008
	(0.179)	(0.206)	(0.205)	(0.185)	(0.183)	(0.181)	(0.200)	(0.167)
Course	−0.180	0.118	0.377*	−0.635***	0.097	0.014	−0.148	−0.183
	(0.175)	(0.206)	(0.202)	(0.206)	(0.187)	(0.184)	(0.216)	(0.174)
Verification	−0.113	0.101	−0.129	−0.212	−0.128	0.163	−0.023	0.088
	(0.157)	(0.166)	(0.166)	(0.181)	(0.186)	(0.189)	(0.185)	(0.177)
C.2 Sample Course Delivery WhatsApp								
Course and verification	0.149	0.386**	−0.088	0.136	0.165	−0.214*	0.035	−0.092
	(0.135)	(0.160)	(0.162)	(0.162)	(0.136)	(0.116)	(0.142)	(0.146)
Course	0.102	0.148	−0.007	0.060	0.390***	−0.302**	−0.190	−0.340**
	(0.134)	(0.151)	(0.161)	(0.168)	(0.142)	(0.118)	(0.148)	(0.164)
Verification	0.013	0.136	0.074	−0.097	0.266*	−0.226*	0.013	−0.346**
	(0.146)	(0.161)	(0.148)	(0.184)	(0.136)	(0.115)	(0.159)	(0.151)
Outcome mean	0	3.505	3.778	2.497	0	1.937	2.33	2.641
Outcome std. dev.	1	1.137	1.113	1.111	1	0.939	1.08	1.041
Outcome range	[-2.6,2.2]	[1,5]	[1,5]	[1,5]	[-3.4,1.7]	[1,5]	[1,5]	[1,5]
Num. LASSO covariates	2	7	1	1	1	7	0	1
Observations	463	463	463	463	463	463	463	463
R^2	0.508	0.521	0.459	0.465	0.451	0.468	0.431	0.470

Notes: We report estimates from OLS regression including randomization block fixed effects. All specifications in all panels include the corresponding outcome variables at baseline as a control. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel C are from a 10-fold cross validation LASSO model with lambda chosen to be that of the minimum average cross-validated error. The treatment indicators, lagged dependent variable, and fixed effects are forced into the LASSO model and covariates are selected from a pool of 72 baseline variables for each specification. Robust standard errors are in parentheses. * denotes p<0.1, ** denotes p<0.05, and *** denotes p<0.01.

Table S17: Treatment effects on mechanisms: Attention to misinformation index's components

	Attention to misinformation index of (1) (1)	How often do you question wether a news is false (2)
Panel A: Treatment effects with covariates in the outcome family		
A.1 Sample course Delivery Online		
Course and verification	0.218	0.177
	(0.173)	(0.140)
Course	0.090	0.073
	(0.171)	(0.139)
Verification	0.226	0.183
	(0.155)	(0.126)
A.2 Sample Course Delivery WhatsApp		
Course and verification	0.297**	0.241**
	(0.127)	(0.103)
Course	0.162	0.131
	(0.134)	(0.108)
Verification	0.114	0.092
	(0.129)	(0.105)
R^2	0.543	0.543
Panel B: IPSW with covariates in the outcome family		
Panel B.1: Sample course Delivery Online		
Course and verification	0.247	0.200
	(0.175)	(0.142)
Course	0.118	0.096
	(0.168)	(0.136)
Verification	0.271*	0.219*
	(0.161)	(0.130)
B.2 Sample Course Delivery WhatsApp		
Course and verification	0.322**	0.261**
	(0.131)	(0.106)
Course	0.178	0.144
	(0.133)	(0.108)
Verification	0.133	0.108
	(0.131)	(0.107)
R^2	0.589	0.589
Panel C: Treatment effects with covariates selected with LASSO		
Panel C.1: Sample course Delivery Online		
Course and verification	0.195	0.158
	(0.172)	(0.139)
Course	0.116	0.094
	(0.172)	(0.140)
Verification	0.220	0.178
	(0.154)	(0.125)
C.2 Sample Course Delivery WhatsApp		
Course and verification	0.267**	0.216**
	(0.129)	(0.105)
Course	0.119	0.096
	(0.143)	(0.115)
Verification	0.099	0.080
	(0.133)	(0.108)
Outcome mean	0	3.901
Outcome std. dev.	1	0.81
Outcome range	[-3.6,1.4]	[1,5]
Num. LASSO covariates	0	0
Observations	463	463
R^2	0.523	0.523

Notes: We report estimates from OLS regression including randomization block fixed effects. All specifications in all panels include the corresponding outcome variables at baseline as a control. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel C are from a 10-fold cross validation LASSO model with lambda chosen to be that of the minimum average cross-validated error. The treatment indicators, lagged dependent variable, and fixed effects are forced into the LASSO model and covariates are selected from a pool of 72 baseline variables for each specification. Robust standard errors are in parentheses. * denotes $p<0.1$, ** denotes $p<0.05$, and *** denotes $p<0.01$.

Table S18: Treatment effects on knowledge to identify information index's components

Panel A: Treatment effects with covariates in the outcome family

	Increase knowledge to identify information index of (1,1,1,-1, -1,-1,1,1,1) (1)	How much knowledge do you have to identify whether a news is false or not? (2)	Knowing how to identify a fake new index (3)	How many recent misinformation cases in Bolivia do you know? (4)	Radio / TV (5)	Newspaper (6)	Internet (7)	WhatsApp (8)	Social media (9)	Conversations (10)
A.1 Sample course Delivery Online										
Course and verification	0.197	0.414***	−0.073	0.190	−0.089*	−0.028	−0.038	−0.146**	0.026	−0.036
	(0.174)	(0.128)	(0.056)	(0.318)	(0.051)	(0.057)	(0.094)	(0.070)	(0.073)	(0.089)
Course	0.245	0.316**	0.012	0.132	−0.029	−0.074	−0.029	−0.035	0.064	−0.044
	(0.193)	(0.140)	(0.062)	(0.275)	(0.064)	(0.065)	(0.102)	(0.066)	(0.065)	(0.096)
Verification	−0.036	−0.083	−0.173***	0.063	−0.036	−0.006	−0.021	0.071	0.030	0.094
	(0.160)	(0.120)	(0.061)	(0.261)	(0.049)	(0.049)	(0.088)	(0.048)	(0.066)	(0.080)
A.2 Sample Course Delivery WhatsApp										
Course and verification	0.076	0.178*	−0.013	0.826***	−0.051	−0.045	0.112	−0.001	−0.065	0.005
	(0.145)	(0.097)	(0.043)	(0.262)	(0.058)	(0.054)	(0.077)	(0.058)	(0.043)	(0.075)
Course	0.072	0.210**	−0.042	0.378	−0.019	0.029	0.029	−0.012	0.047	−0.012
	(0.144)	(0.087)	(0.037)	(0.282)	(0.052)	(0.051)	(0.080)	(0.051)	(0.040)	(0.076)
Verification	0.025	−0.102	0.012	0.374	0.020	0.009	−0.040	−0.012	−0.091*	0.024
	(0.165)	(0.110)	(0.041)	(0.290)	(0.065)	(0.052)	(0.077)	(0.055)	(0.054)	(0.081)
R^2	0.450	0.575	0.474	0.454	0.347	0.380	0.338	0.484	0.391	0.414

Panel B: IPSW with covariates in the outcome family

Panel B.1: Sample course Delivery Online										
Course and verification	0.186	0.408***	−0.085	0.213	−0.086	−0.025	−0.060	−0.169**	0.021	−0.057
	(0.172)	(0.129)	(0.055)	(0.305)	(0.053)	(0.056)	(0.097)	(0.071)	(0.071)	(0.091)
Course	0.228	0.370***	−0.016	0.188	−0.035	−0.071	−0.004	−0.033	0.082	−0.051
	(0.190)	(0.136)	(0.061)	(0.257)	(0.061)	(0.060)	(0.099)	(0.067)	(0.060)	(0.099)
Verification	−0.008	−0.068	−0.175***	0.115	−0.050	0.003	−0.036	0.067	0.043	0.072
	(0.165)	(0.120)	(0.059)	(0.262)	(0.051)	(0.049)	(0.091)	(0.053)	(0.065)	(0.086)
B.2 Sample Course Delivery WhatsApp										
Course and verification	0.068	0.179*	−0.025	0.796***	−0.048	−0.023	0.103	−0.009	−0.070	0.014
	(0.147)	(0.101)	(0.043)	(0.265)	(0.062)	(0.058)	(0.077)	(0.059)	(0.044)	(0.076)
Course	0.065	0.219**	−0.047	0.405	−0.022	0.036	0.029	−0.018	0.040	−0.016
	(0.147)	(0.089)	(0.037)	(0.285)	(0.058)	(0.053)	(0.081)	(0.054)	(0.042)	(0.078)
Verification	−0.015	−0.073	0.012	0.277	0.054	0.040	−0.055	−0.025	−0.101*	0.010
	(0.158)	(0.109)	(0.038)	(0.292)	(0.069)	(0.055)	(0.075)	(0.056)	(0.054)	(0.081)
R^2	0.487	0.590	0.518	0.503	0.379	0.405	0.378	0.531	0.443	0.442

Panel C: Treatment effects with covariates selected with LASSO

Panel C.1: Sample course Delivery Online										
Course and verification	0.212	0.369***	−0.070	0.010	−0.136**	−0.033	−0.048	−0.148**	0.001	−0.036
	(0.161)	(0.115)	(0.056)	(0.282)	(0.054)	(0.055)	(0.086)	(0.070)	(0.069)	(0.089)
Course	0.374**	0.388***	0.024	0.076	−0.077	−0.089	−0.108	−0.042	0.050	−0.039
	(0.183)	(0.120)	(0.063)	(0.280)	(0.058)	(0.062)	(0.092)	(0.066)	(0.063)	(0.094)
Verification	0.039	−0.036	−0.157**	0.038	−0.050	−0.036	−0.062	0.074	0.050	0.122
	(0.148)	(0.092)	(0.061)	(0.254)	(0.048)	(0.048)	(0.079)	(0.050)	(0.063)	(0.079)
C.2 Sample Course Delivery WhatsApp										
Course and verification	0.092	0.198**	−0.011	0.652***	−0.089	−0.051	0.080	−0.006	−0.082*	0.012
	(0.137)	(0.098)	(0.041)	(0.247)	(0.059)	(0.057)	(0.070)	(0.058)	(0.041)	(0.074)
Course	0.129	0.291***	−0.036	0.220	−0.047	0.016	−0.000	−0.004	0.047	0.000
	(0.128)	(0.086)	(0.036)	(0.235)	(0.053)	(0.053)	(0.073)	(0.051)	(0.038)	(0.074)
Verification	0.080	−0.036	0.020	0.532*	0.007	−0.014	−0.071	−0.005	−0.078	0.042
	(0.152)	(0.105)	(0.040)	(0.274)	(0.062)	(0.055)	(0.073)	(0.053)	(0.049)	(0.078)
Outcome mean	0	3.311	0.406	1.678	0.114	0.117	0.594	0.76	0.868	0.49
Outcome std. dev.	1	0.818	0.318	1.781	0.319	0.321	0.492	0.373	0.339	0.5
Outcome range	[-3.3,2.5]	[1,5]	[-0.8,1]	[0,5]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]	[0,1]
Num. LASSO covariates	8	15	1	19	3	1	15	1	7	1
Observations	463	463	463	463	463	463	463	463	463	463
R^2	0.537	0.663	0.478	0.574	0.381	0.382	0.444	0.484	0.444	0.423

Notes: We report estimates from OLS regression including randomization block fixed effects. All specifications in all panels include the corresponding outcome variables at baseline as a control. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel C are from a 10-fold cross validation LASSO model with lambda chosen to be that of the minimum average cross-validated error. The treatment indicators, lagged dependent variable, and fixed effects are forced into the LASSO model and covariates are selected from a pool of 72 baseline variables for each specification. Robust standard errors are in parentheses. * denotes p<0.1, ** denotes p<0.05, and *** denotes p<0.01.

Table S19: Treatment effects on knowledge to verify information index's components

	Increase knowledge to verify information index of (1,1,1) (1)	How much knowledge do you have to verify if a doubtful news is false or not? (2)	Main ways to verify news index (3)	How many verifiers or fact-checkers do you know? (4)
Panel A: Treatment effects with covariates in the outcome family				
A.1 Sample course Delivery Online				
Course and verification	0.606***	0.484***	0.127**	0.390**
	(0.133)	(0.126)	(0.055)	(0.165)
Course	0.569***	0.391***	0.094	0.535***
	(0.160)	(0.133)	(0.064)	(0.185)
Verification	−0.004	−0.012	0.037	−0.140
	(0.117)	(0.112)	(0.050)	(0.130)
A.2 Sample Course Delivery WhatsApp				
Course and verification	0.154	0.125	0.034	0.071
	(0.128)	(0.099)	(0.054)	(0.149)
Course	0.215*	0.283***	0.135***	−0.372**
	(0.120)	(0.089)	(0.047)	(0.154)
Verification	0.087	0.015	0.040	0.012
	(0.133)	(0.099)	(0.048)	(0.174)
R^2	0.604	0.637	0.468	0.554
Panel B: IPSW with covariates in the outcome family				
Panel B.1: Sample course Delivery Online				
Course and verification	0.624***	0.497***	0.131**	0.410**
	(0.138)	(0.130)	(0.058)	(0.169)
Course	0.643***	0.478***	0.110*	0.556***
	(0.160)	(0.133)	(0.066)	(0.181)
Verification	0.028	0.002	0.048	−0.117
	(0.124)	(0.113)	(0.053)	(0.142)
B.2 Sample Course Delivery WhatsApp				
Course and verification	0.109	0.119	0.018	0.018
	(0.128)	(0.106)	(0.058)	(0.150)
Course	0.221*	0.276***	0.141***	−0.377**
	(0.120)	(0.093)	(0.048)	(0.158)
Verification	0.064	0.004	0.053	−0.086
	(0.128)	(0.102)	(0.050)	(0.173)
R^2	0.620	0.656	0.477	0.569
Panel C: Treatment effects with covariates selected with LASSO				
Panel C.1: Sample course Delivery Online				
Course and verification	0.573***	0.424***	0.138***	0.388**
	(0.135)	(0.114)	(0.052)	(0.157)
Course	0.634***	0.440***	0.123**	0.551***
	(0.140)	(0.120)	(0.055)	(0.180)
Verification	−0.002	−0.034	0.055	−0.119
	(0.111)	(0.101)	(0.045)	(0.130)
C.2 Sample Course Delivery WhatsApp				
Course and verification	0.193	0.139	0.049	0.159
	(0.120)	(0.098)	(0.050)	(0.152)
Course	0.235**	0.306***	0.127***	−0.345**
	(0.116)	(0.087)	(0.047)	(0.155)
Verification	0.045	0.024	0.014	0.035
	(0.127)	(0.102)	(0.050)	(0.181)
Outcome mean	0	3.305	0.642	1.369
Outcome std. dev.	1	0.848	0.343	1.122
Outcome range	[-3.3,2.1]	[1,5]	[-0.3,1]	[0,3]
Num. LASSO covariates	7	14	17	1
Observations	463	463	463	463
R^2	0.656	0.682	0.538	0.556

Notes: We report estimates from OLS regression including randomization block fixed effects. All specifications in all panels include the corresponding outcome variables at baseline as a control. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel C are from a 10-fold cross validation LASSO model with lambda chosen to be that of the minimum average cross-validated error. The treatment indicators, lagged dependent variable, and fixed effects are forced into the LASSO model and covariates are selected from a pool of 72 baseline variables for each specification. Robust standard errors are in parentheses. * denotes p<0.1, ** denotes p<0.05, and *** denotes p<0.01.

Table S20: Treatment effects on consumption behavior index's components

Panel A: Treatment effects with covariates in the outcome family

	Increase consumption behavior of traditional sources index of (1,1,1)	Radio / TV	Newspaper	Internet	Decrease consumption behavior of social media sources index of (-1,-1,-1)	WhatsApp	Social media	Conversations
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A.1 Sample course Delivery Online								
Course and verification	0.054	−0.020	0.048	0.132	−0.228	0.650**	−0.101	0.355
	(0.154)	(0.222)	(0.240)	(0.236)	(0.165)	(0.292)	(0.172)	(0.220)
Course	−0.148	0.062	−0.029	−0.525*	−0.088	0.364	−0.088	0.103
	(0.177)	(0.217)	(0.260)	(0.293)	(0.157)	(0.277)	(0.186)	(0.226)
Verification	0.143	0.273	0.181	0.072	−0.258*	0.488*	0.104	0.256
	(0.130)	(0.189)	(0.229)	(0.214)	(0.151)	(0.270)	(0.154)	(0.191)
A.2 Sample Course Delivery WhatsApp								
Course and verification	−0.191	−0.213	−0.157	−0.278	0.269*	−0.364*	−0.295*	−0.166
	(0.159)	(0.238)	(0.209)	(0.226)	(0.145)	(0.191)	(0.171)	(0.215)
Course	−0.081	0.338	−0.034	−0.554**	0.364***	−0.456**	−0.200	−0.506***
	(0.171)	(0.251)	(0.231)	(0.232)	(0.120)	(0.184)	(0.162)	(0.186)
Verification	−0.069	0.092	0.034	−0.292	0.464***	−0.448**	−0.220	−0.795***
	(0.154)	(0.222)	(0.223)	(0.266)	(0.138)	(0.195)	(0.184)	(0.191)
R^2	0.457	0.489	0.455	0.439	0.531	0.527	0.506	0.471

Panel B: IPSW with covariates in the outcome family

Panel B.1: Sample course Delivery Online								
Course and verification	0.059	0.052	0.012	0.115	−0.218	0.649**	−0.106	0.313
	(0.160)	(0.223)	(0.247)	(0.249)	(0.170)	(0.298)	(0.179)	(0.217)
Course	−0.199	0.001	−0.114	−0.545*	−0.055	0.350	−0.113	0.035
	(0.182)	(0.217)	(0.282)	(0.283)	(0.158)	(0.282)	(0.191)	(0.208)
Verification	0.146	0.334*	0.123	0.081	−0.264*	0.503*	0.101	0.262
	(0.137)	(0.197)	(0.234)	(0.227)	(0.155)	(0.273)	(0.162)	(0.198)
B.2 Sample Course Delivery WhatsApp								
Course and verification	−0.169	−0.127	−0.206	−0.259	0.303**	−0.387**	−0.302*	−0.231
	(0.162)	(0.242)	(0.220)	(0.229)	(0.141)	(0.184)	(0.169)	(0.209)
Course	−0.041	0.443*	−0.063	−0.520**	0.388***	−0.432**	−0.219	−0.552***
	(0.171)	(0.251)	(0.238)	(0.231)	(0.120)	(0.184)	(0.164)	(0.182)
Verification	−0.033	0.203	−0.042	−0.225	0.489***	−0.510**	−0.213	−0.833***
	(0.165)	(0.243)	(0.248)	(0.261)	(0.140)	(0.203)	(0.185)	(0.185)
R^2	0.484	0.523	0.485	0.453	0.561	0.563	0.534	0.509

Panel C: Treatment effects with covariates selected with LASSO

Panel C.1: Sample course Delivery Online								
Course and verification	0.115	0.050	0.053	0.157	−0.265	0.642**	−0.142	0.365*
	(0.152)	(0.223)	(0.241)	(0.236)	(0.167)	(0.286)	(0.176)	(0.214)
Course	−0.116	0.189	−0.024	−0.582**	−0.072	0.254	−0.160	0.058
	(0.164)	(0.204)	(0.271)	(0.294)	(0.161)	(0.283)	(0.194)	(0.207)
Verification	0.162	0.265	0.169	0.079	−0.264*	0.444*	0.111	0.310*
	(0.124)	(0.176)	(0.228)	(0.217)	(0.147)	(0.269)	(0.161)	(0.185)
C.2 Sample Course Delivery WhatsApp								
Course and verification	−0.107	−0.075	−0.217	−0.232	0.255*	−0.414**	−0.273	−0.166
	(0.150)	(0.233)	(0.208)	(0.219)	(0.144)	(0.194)	(0.172)	(0.213)
Course	−0.082	0.334	−0.079	−0.563**	0.404***	−0.476***	−0.238	−0.563***
	(0.161)	(0.224)	(0.235)	(0.222)	(0.117)	(0.176)	(0.158)	(0.184)
Verification	−0.030	0.052	0.034	−0.188	0.443***	−0.452**	−0.241	−0.738***
	(0.151)	(0.215)	(0.222)	(0.259)	(0.137)	(0.192)	(0.190)	(0.181)
Outcome mean	0	4.175	4.108	3.844	0	3.609	5.037	3.86
Outcome std. dev.	1	1.484	1.445	1.51	1	1.526	1.168	1.327
Outcome range	[-2.8,1.8]	[1,6]	[1,6]	[1,6]	[-1.7,3.3]	[1,6]	[1,6]	[1,6]
Num. LASSO covariates	12	20	0	7	3	1	0	9
Observations	463	463	463	463	463	463	463	463
R^2	0.515	0.565	0.443	0.466	0.538	0.511	0.487	0.515

Notes: We report estimates from OLS regression including randomization block fixed effects. All specifications in all panels include the corresponding outcome variables at baseline as a control. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel C are from a 10-fold cross validation LASSO model with lambda chosen to be that of the minimum average cross-validated error. The treatment indicators, lagged dependent variable, and fixed effects are forced into the LASSO model and covariates are selected from a pool of 72 baseline variables for each specification. Robust standard errors are in parentheses. * denotes p<0.1, ** denotes p<0.05, and *** denotes p<0.01.

Table S21: Treatment effects on sharing behavior index's components

Panel A: Treatment effects with covariates in the outcome family

	Decrease sharing behavior index of (1,-1, -1,-1) (1)	If you know a story is false, how often do you share its falsehood with others on social media? (2)	How often do you share news you receive on WhatsApp and social media (3)	Share scenario 1 withoout verification† (4)	Share scenario 3 withoout verification†† (5)
A.1 Sample course Delivery Online					
Course and verification	0.152	0.180	0.082	0.006	−0.382*
	(0.184)	(0.212)	(0.148)	(0.185)	(0.197)
Course	0.292	−0.092	−0.064	−0.250	−0.378*
	(0.232)	(0.218)	(0.248)	(0.192)	(0.211)
Verification	0.374**	0.098	−0.156	−0.239	−0.365*
	(0.184)	(0.206)	(0.175)	(0.178)	(0.189)
A.2 Sample Course Delivery WhatsApp					
Course and verification	−0.023	0.212	0.171	0.001	−0.031
	(0.143)	(0.200)	(0.140)	(0.146)	(0.175)
Course	0.102	0.094	0.114	−0.128	−0.240
	(0.144)	(0.202)	(0.144)	(0.145)	(0.159)
Verification	0.243	0.198	−0.120	0.021	−0.296
	(0.169)	(0.221)	(0.151)	(0.161)	(0.192)
R^2	0.381	0.444	0.401	0.428	0.456

Panel B: IPSW with covariates in the outcome family

Panel B.1: Sample course Delivery Online					
Course and verification	0.142	0.190	0.161	0.004	−0.477**
	(0.179)	(0.214)	(0.221)	(0.185)	(0.215)
Course	0.366*	−0.091	−0.202	−0.201	−0.488**
	(0.198)	(0.218)	(0.247)	(0.182)	(0.223)
Verification	0.361**	0.100	−0.066	−0.234	−0.508**
	(0.180)	(0.213)	(0.211)	(0.175)	(0.212)
B.2 Sample Course Delivery WhatsApp					
Course and verification	−0.018	0.154	0.100	0.034	0.013
	(0.143)	(0.207)	(0.139)	(0.146)	(0.177)
Course	0.082	0.009	0.076	−0.097	−0.230
	(0.143)	(0.204)	(0.144)	(0.143)	(0.167)
Verification	0.278*	0.257	−0.150	0.085	−0.371*
	(0.163)	(0.226)	(0.150)	(0.160)	(0.200)
R^2	0.437	0.469	0.411	0.446	0.473

Panel C: Treatment effects with covariates selected with LASSO

Panel C.1: Sample course Delivery Online					
Course and verification	0.118	0.141	0.194	−0.019	−0.492**
	(0.172)	(0.195)	(0.127)	(0.190)	(0.198)
Course	0.331	−0.068	0.032	−0.235	−0.591***
	(0.215)	(0.214)	(0.220)	(0.201)	(0.198)
Verification	0.423**	0.099	−0.044	−0.279	−0.518***
	(0.169)	(0.195)	(0.156)	(0.180)	(0.177)
C.2 Sample Course Delivery WhatsApp					
Course and verification	−0.022	0.217	0.133	0.001	−0.005
	(0.138)	(0.204)	(0.138)	(0.146)	(0.164)
Course	0.131	0.074	0.037	−0.087	−0.268*
	(0.146)	(0.204)	(0.138)	(0.150)	(0.151)
Verification	0.319*	0.221	−0.289*	0.026	−0.272
	(0.170)	(0.224)	(0.154)	(0.164)	(0.185)
Outcome mean	0	2.775	0	1.756	1.955
Outcome std. dev.	1	1.307	0.999	0.982	1.168
Outcome range	[-3,5.1]	[1,6]	[-7,1.6]	[1,5]	[1,5]
Num. LASSO covariates	6	1	6	0	11
Observations	463	463	463	463	463
R^2	0.441	0.454	0.492	0.406	0.528

Notes: We report estimates from OLS regression including randomization block fixed effects. All specifications in all panels include the corresponding outcome variables at baseline as a control. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel C are from a 10-fold cross validation LASSO model with lambda chosen to be that of the minimum average cross-validated error. The treatment indicators, lagged dependent variable, and fixed effects are forced into the LASSO model and covariates are selected from a pool of 72 baseline variables for each specification. Robust standard errors are in parentheses. †: imagine you received an that came to you on WhatsApp with a screenshot of the Twitter account of a well-known person saying something very controversial. ††: imagine you received an audio that came to you by WhatsApp reporting evidence of corruption of a politician that you already suspected and you did not like. * denotes p<0.1, ** denotes p<0.05, and *** denotes p<0.01.

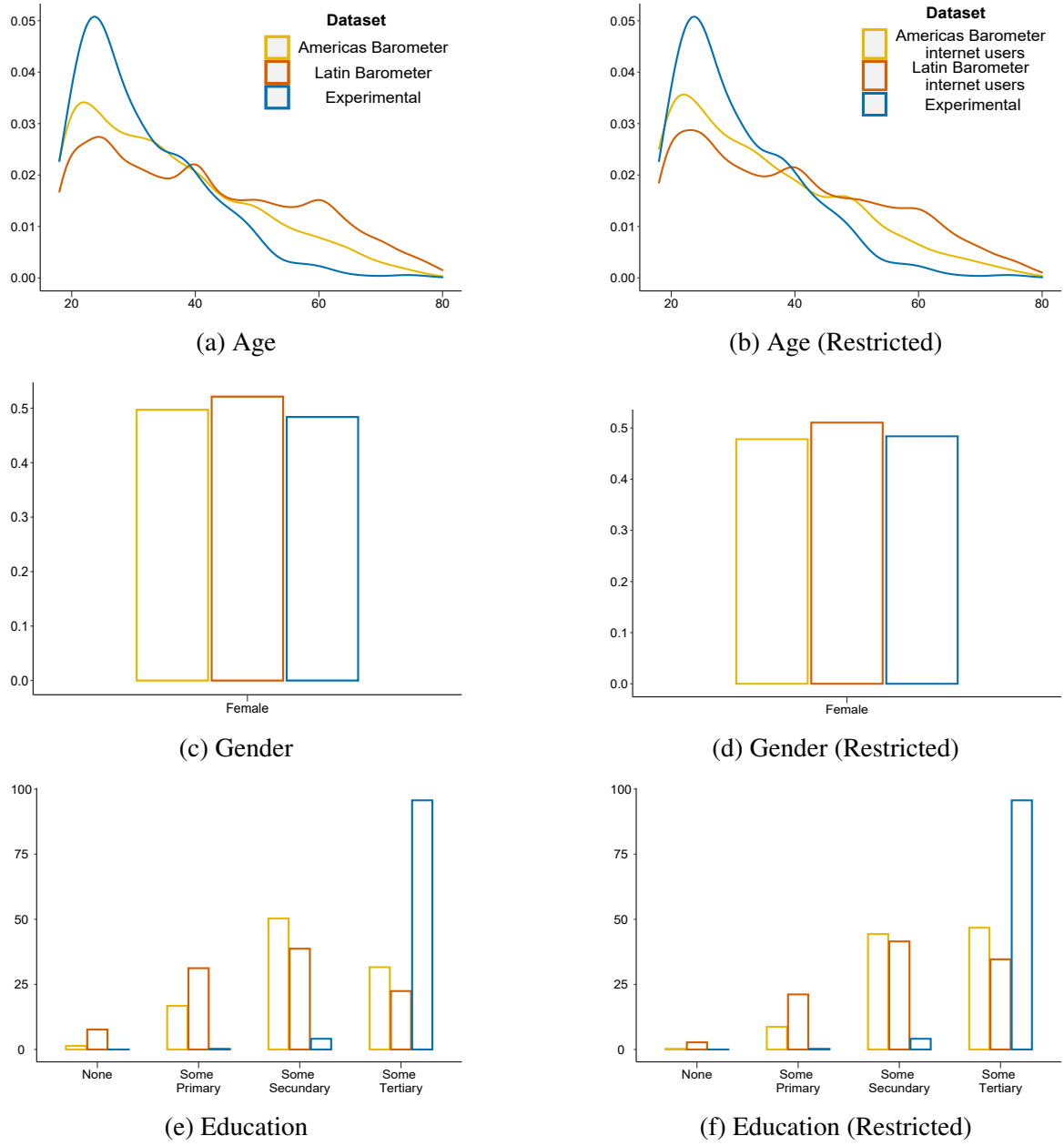
Table S22: Treatment effects on verifying behavior index's components

	Increase verifying behavior index of (1) (1)	How often do you verify news that you doubt may be false before sharing it? (2)
Panel A: Treatment effects with covariates in the outcome family		
A.1 Sample course Delivery Online		
Course and verification	0.135	0.126
	(0.175)	(0.163)
Course	0.190	0.177
	(0.197)	(0.183)
Verification	−0.012	−0.011
	(0.179)	(0.166)
A.2 Sample Course Delivery WhatsApp		
Course and verification	0.038	0.035
	(0.144)	(0.134)
Course	0.049	0.045
	(0.152)	(0.141)
Verification	0.208	0.194
	(0.131)	(0.122)
R^2	0.437	0.437
Panel B: IPSW with covariates in the outcome family		
Panel B.1: Sample course Delivery Online		
Course and verification	0.151	0.140
	(0.174)	(0.162)
Course	0.259	0.241
	(0.189)	(0.176)
Verification	−0.007	−0.007
	(0.181)	(0.168)
B.2 Sample Course Delivery WhatsApp		
Course and verification	0.025	0.024
	(0.147)	(0.137)
Course	0.038	0.035
	(0.158)	(0.147)
Verification	0.189	0.176
	(0.144)	(0.134)
R^2	0.499	0.499
Panel C: Treatment effects with covariates selected with LASSO		
Panel C.1: Sample course Delivery Online		
Course and verification	0.064	0.060
	(0.157)	(0.146)
Course	0.180	0.168
	(0.181)	(0.168)
Verification	−0.069	−0.064
	(0.164)	(0.153)
C.2 Sample Course Delivery WhatsApp		
Course and verification	0.045	0.041
	(0.139)	(0.129)
Course	0.100	0.093
	(0.142)	(0.132)
Verification	0.194	0.180
	(0.128)	(0.119)
Outcome mean	0	3.808
Outcome std. dev.	1	0.93
Outcome range	[-3,1.3]	[1,5]
Num. LASSO covariates	7	7
Observations	463	463
R^2	0.506	0.506

Notes: We report estimates from OLS regression including randomization block fixed effects. All specifications in all panels include the corresponding outcome variables at baseline as a control. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel C are from a 10-fold cross validation LASSO model with lambda chosen to be that of the minimum average cross-validated error. The treatment indicators, lagged dependent variable, and fixed effects are forced into the LASSO model and covariates are selected from a pool of 72 baseline variables for each specification. Robust standard errors are in parentheses. * denotes $p<0.1$, ** denotes $p<0.05$, and *** denotes $p<0.01$.

Sample Comparison

Figure S6: Comparison of demographics in the restricted and unrestricted Americas Barometer sample and the Latin Barometer sample to internet users, and experimental sample respondents



(a) Age

(b) Age (Restricted)

(c) Gender

(d) Gender (Restricted)

(e) Education

(f) Education (Restricted)

Notes: Additional summary statistic comparisons are in Table S23.

Table S23: Summary statistics of comparable demographics in the restricted and unrestricted Americas Barometer sample and the Latin Barometer sample to internet users, and the experimental sample

	Americas Barometer sample	Latin Barometer sample	Americas Barometer internet users sample	Latin Barometer internet users sample	Experimental sample
Age	38.575	40.962	38.165	39.780	31.114
	15.002	16.436	15.023	15.993	10.203
	3002	20077	1393	10163	463
Gender	0.497	0.521	0.478	0.511	0.484
	0.500	0.500	0.500	0.500	0.500
	3002	20077	1393	10163	463
Education	2.121	1.759	2.376	2.080	2.955
	0.723	0.886	0.650	0.813	0.218
	3002	20077	1393	10163	463

Notes: For every variable, each row shows the mean, standard deviation, and number of observations. The Americas Barometer sample is for the year 2021 and the Latin Barometer sample for the year 2020. The Americas Baromoter restricted to internet users and the Latin Barometer restricted to internet users are the subset of respondents that have connection or a contract of internet at home.

Table S24: Treatment effects on indexes by course delivered online and via WhatsApp with post-stratifications weights to mimic the Americas Barometer sample restricted to internet users

	Mis-information importance index (1)	Decrease likelihood of false of traditional sources index (2)	Increase likelihood of false of social media sources index (3)	Trust traditional sources index (4)	Disrust social media sources index (5)	Attention to mis-information index (6)	Increase knowledge to identify information index (7)	Increase knowledge to verify information index (8)	Increase consumption behavior of traditional sources index (9)	Decrease consumption behavior of social media sources index (10)	Decrease sharing behavior index (11)	Increase verifying behavior index (12)
Panel A: Treatment effects with covariates in the outcome family												
A.1 Sample course Delivery Online												
Course and verification	−0.018	0.328*	−0.112	−0.009	−0.049	0.224	0.213	0.599***	0.056	−0.211	0.133	0.166
	(0.184)	(0.184)	(0.156)	(0.181)	(0.187)	(0.174)	(0.176)	(0.134)	(0.155)	(0.165)	(0.195)	(0.176)
Course	0.116	−0.069	0.188	−0.203	0.054	0.097	0.296	0.582***	−0.178	−0.115	0.281	0.216
	(0.215)	(0.195)	(0.176)	(0.180)	(0.185)	(0.174)	(0.193)	(0.160)	(0.178)	(0.155)	(0.236)	(0.197)
Verification	−0.070	0.029	−0.176	−0.119	−0.059	0.265*	−0.005	−0.008	0.136	−0.247	0.376**	0.011
	(0.164)	(0.159)	(0.168)	(0.165)	(0.191)	(0.159)	(0.164)	(0.115)	(0.134)	(0.152)	(0.189)	(0.179)
A.2 Sample Course Delivery WhatsApp												
Course and verification	−0.017	0.085	0.347**	0.109	0.099	0.309**	0.070	0.173	−0.197	0.278*	−0.016	0.077
	(0.146)	(0.144)	(0.163)	(0.141)	(0.145)	(0.142)	(0.148)	(0.133)	(0.159)	(0.144)	(0.142)	(0.154)
Course	−0.247	−0.147	0.169	0.102	0.268*	0.101	0.058	0.150	−0.060	0.359***	0.079	−0.013
	(0.160)	(0.141)	(0.166)	(0.142)	(0.154)	(0.149)	(0.143)	(0.123)	(0.174)	(0.127)	(0.144)	(0.156)
Verification	−0.180	0.066	0.233	0.029	0.185	0.077	0.026	0.072	−0.052	0.467***	0.219	0.191
	(0.163)	(0.149)	(0.141)	(0.152)	(0.140)	(0.136)	(0.167)	(0.133)	(0.158)	(0.140)	(0.168)	(0.131)
R^2	0.430	0.766	0.870	0.700	0.804	0.883	0.507	0.822	0.483	0.582	0.573	0.704
Panel B: IPSW with covariates in the outcome family												
Panel B.1: Sample course Delivery Online												
Course and verification	−0.010	0.310*	−0.106	−0.001	−0.038	0.257	0.198	0.619***	0.057	−0.202	0.145	0.185
	(0.184)	(0.181)	(0.157)	(0.181)	(0.183)	(0.177)	(0.174)	(0.139)	(0.160)	(0.169)	(0.181)	(0.176)
Course	0.140	−0.027	0.194	−0.194	0.080	0.136	0.277	0.659***	−0.237	−0.083	0.326	0.295
	(0.222)	(0.197)	(0.181)	(0.177)	(0.178)	(0.170)	(0.192)	(0.162)	(0.185)	(0.157)	(0.200)	(0.190)
Verification	−0.107	0.008	−0.133	−0.145	−0.022	0.310*	0.023	0.027	0.136	−0.255	0.357**	0.019
	(0.169)	(0.166)	(0.175)	(0.172)	(0.191)	(0.163)	(0.169)	(0.124)	(0.141)	(0.157)	(0.181)	(0.182)
B.2 Sample Course Delivery WhatsApp												
Course and verification	−0.082	0.105	0.328**	0.109	0.110	0.336**	0.064	0.128	−0.175	0.313**	−0.017	0.071
	(0.164)	(0.148)	(0.163)	(0.142)	(0.146)	(0.141)	(0.150)	(0.132)	(0.162)	(0.140)	(0.141)	(0.155)
Course	−0.315*	−0.151	0.174	0.104	0.298**	0.138	0.055	0.164	−0.022	0.389***	0.072	−0.004
	(0.170)	(0.151)	(0.158)	(0.143)	(0.150)	(0.142)	(0.146)	(0.122)	(0.173)	(0.124)	(0.141)	(0.158)
Verification	−0.163	0.129	0.236*	0.064	0.170	0.105	−0.015	0.046	−0.018	0.491***	0.253	0.181
	(0.174)	(0.146)	(0.140)	(0.149)	(0.138)	(0.136)	(0.160)	(0.128)	(0.168)	(0.141)	(0.161)	(0.142)
R^2	0.430	0.781	0.878	0.725	0.819	0.895	0.537	0.830	0.503	0.613	0.624	0.725
Panel C: Treatment effects with covariates selected with LASSO												
Panel C.1: Sample course Delivery Online												
Course and verification	−0.046	0.444***	−0.191	0.060	−0.002	0.203	0.225	0.554***	0.124	−0.250	0.110	0.078
	(0.186)	(0.172)	(0.160)	(0.185)	(0.186)	(0.170)	(0.166)	(0.137)	(0.151)	(0.165)	(0.179)	(0.161)
Course	0.109	−0.024	0.236	−0.183	0.081	0.135	0.421**	0.640***	−0.122	−0.105	0.314	0.189
	(0.209)	(0.192)	(0.185)	(0.178)	(0.187)	(0.173)	(0.182)	(0.138)	(0.161)	(0.158)	(0.218)	(0.181)
Verification	−0.072	−0.008	−0.192	−0.105	−0.117	0.262*	0.071	−0.011	0.148	−0.265*	0.418**	−0.070
	(0.165)	(0.151)	(0.162)	(0.159)	(0.188)	(0.152)	(0.155)	(0.111)	(0.124)	(0.148)	(0.170)	(0.168)
C.2 Sample Course Delivery WhatsApp												
Course and verification	−0.048	0.071	0.320**	0.161	0.157	0.294*	0.092	0.195	−0.108	0.267*	−0.019	0.088
	(0.146)	(0.132)	(0.148)	(0.141)	(0.145)	(0.167)	(0.139)	(0.122)	(0.155)	(0.144)	(0.138)	(0.147)
Course	−0.272*	−0.147	0.137	0.191	0.310**	−0.011	0.117	0.165	−0.012	0.373***	0.106	0.033
	(0.159)	(0.139)	(0.161)	(0.153)	(0.148)	(0.197)	(0.126)	(0.118)	(0.182)	(0.139)	(0.144)	(0.151)
Verification	−0.187	0.051	0.241*	0.046	0.226	0.051	0.082	0.028	−0.017	0.438***	0.304*	0.175
	(0.166)	(0.138)	(0.131)	(0.150)	(0.142)	(0.145)	(0.153)	(0.126)	(0.154)	(0.140)	(0.170)	(0.128)
Outcome mean	0	0	0	0	0	0	0	0	0	0	0	0
Outcome std. dev.	1	1	1	1	1	1	1	1	1	1	1	1
Outcome range	[-3.2,0.6]	[-4.1,3.1]	[-3.8,2.5]	[-2.6,2.2]	[-3.4,1.7]	[-3.6,1.4]	[-3.3,2.5]	[-3.3,2.1]	[-2.8,1.8]	[-1.7,3.3]	[-3,5.1]	[-3,1.3]
Num. LASSO covariates	0	9	8	2	1	0	8	7	12	3	6	7
Observations	463	463	463	463	463	463	463	463	463	463	463	463
R^2	0.419	0.795	0.875	0.706	0.802	0.872	0.584	0.846	0.534	0.587	0.615	0.738

Notes: We report estimates from OLS regression including randomization block fixed effects. All specifications in all panels include the corresponding outcome variables at baseline as a control. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel C are from a 10-fold cross validation LASSO model with lambda chosen to be that of the minimum average cross-validated error. The treatment indicators, lagged dependent variable, and fixed effects are forced into the LASSO model and covariates are selected from a pool of 72 baseline variables for each specification. Robust standard errors are in parentheses. * denotes p<0.1, ** denotes p<0.05, and *** denotes p<0.01.

Table S25: Treatment effects on indexes by course delivered online and via WhatsApp with post-stratifications weights to mimic the Latin Barometer sample restricted to internet users

Panel A: Treatment effects with covariates in the outcome family

	Mis-information importance index (1)	Decrease likelihood of false of traditional sources index (2)	Increase likelihood of false of social media sources index (3)	Trust traditional sources index (4)	Disrust social media sources index (5)	Attention to mis-information index (6)	Increase knowledge to identify information index (7)	Increase knowledge to verify information index (8)	Increase consumption behavior of traditional sources index (9)	Decrease consumption behavior of social media sources index (10)	Decrease sharing behavior index (11)	Increase verifying behavior index (12)
A.1 Sample course Delivery Online												
Course and verification	−0.046	0.412**	−0.089	0.039	−0.051	0.240	0.255	0.560***	0.063	−0.182	0.074	0.211
	(0.177)	(0.202)	(0.170)	(0.195)	(0.196)	(0.180)	(0.191)	(0.141)	(0.161)	(0.167)	(0.231)	(0.189)
Course	0.024	−0.049	0.282	−0.235	0.032	0.157	0.421**	0.606***	−0.253	−0.184	0.228	0.268
	(0.238)	(0.199)	(0.196)	(0.189)	(0.188)	(0.179)	(0.198)	(0.163)	(0.181)	(0.152)	(0.247)	(0.199)
Verification	−0.142	0.126	−0.141	−0.106	−0.013	0.346**	0.076	−0.033	0.134	−0.237	0.362*	0.049
	(0.161)	(0.178)	(0.183)	(0.168)	(0.203)	(0.166)	(0.178)	(0.117)	(0.145)	(0.162)	(0.210)	(0.190)
A.2 Sample Course Delivery WhatsApp												
Course and verification	−0.004	0.107	0.321*	0.137	0.060	0.333**	0.065	0.199	−0.189	0.288**	0.002	0.139
	(0.147)	(0.144)	(0.182)	(0.145)	(0.155)	(0.148)	(0.160)	(0.137)	(0.159)	(0.144)	(0.143)	(0.159)
Course	−0.205	−0.131	0.169	0.119	0.226	0.143	0.020	0.087	−0.103	0.409***	0.046	−0.031
	(0.156)	(0.135)	(0.175)	(0.140)	(0.157)	(0.154)	(0.144)	(0.129)	(0.175)	(0.125)	(0.147)	(0.157)
Verification	−0.174	0.065	0.173	0.059	0.117	0.042	0.042	0.081	−0.032	0.489***	0.170	0.186
	(0.156)	(0.157)	(0.153)	(0.157)	(0.152)	(0.139)	(0.176)	(0.135)	(0.163)	(0.144)	(0.168)	(0.125)
R^2	0.495	0.856	0.926	0.797	0.883	0.936	0.571	0.890	0.541	0.642	0.690	0.812

Panel B: IPSW with covariates in the outcome family

Panel B.1: Sample course Delivery Online												
Course and verification	0.002	0.275	−0.115	−0.026	−0.038	0.247	0.186	0.624***	0.059	−0.218	0.142	0.151
	(0.186)	(0.179)	(0.155)	(0.178)	(0.182)	(0.175)	(0.172)	(0.138)	(0.160)	(0.170)	(0.179)	(0.174)
Course	0.174	−0.036	0.175	−0.185	0.093	0.118	0.228	0.643***	−0.199	−0.055	0.366*	0.259
	(0.211)	(0.197)	(0.175)	(0.175)	(0.181)	(0.168)	(0.190)	(0.160)	(0.182)	(0.158)	(0.198)	(0.189)
Verification	−0.076	−0.029	−0.149	−0.148	−0.048	0.271*	−0.008	0.028	0.146	−0.264*	0.361**	−0.007
	(0.170)	(0.161)	(0.170)	(0.171)	(0.189)	(0.161)	(0.165)	(0.124)	(0.137)	(0.155)	(0.180)	(0.181)
B.2 Sample Course Delivery WhatsApp												
Course and verification	−0.084	0.096	0.332**	0.096	0.112	0.322**	0.068	0.109	−0.169	0.303**	−0.018	0.025
	(0.162)	(0.150)	(0.143)	(0.140)	(0.140)	(0.131)	(0.147)	(0.128)	(0.162)	(0.141)	(0.143)	(0.147)
Course	−0.343**	−0.185	0.245	0.063	0.358**	0.178	0.065	0.221*	−0.041	0.388***	0.082	0.038
	(0.174)	(0.155)	(0.148)	(0.143)	(0.149)	(0.133)	(0.147)	(0.120)	(0.171)	(0.120)	(0.143)	(0.158)
Verification	−0.163	0.117	0.275**	0.042	0.199	0.133	−0.015	0.064	−0.033	0.489***	0.278*	0.189
	(0.176)	(0.146)	(0.132)	(0.149)	(0.136)	(0.131)	(0.158)	(0.128)	(0.165)	(0.140)	(0.163)	(0.144)
R^2	0.385	0.499	0.519	0.533	0.485	0.589	0.487	0.620	0.484	0.561	0.437	0.499

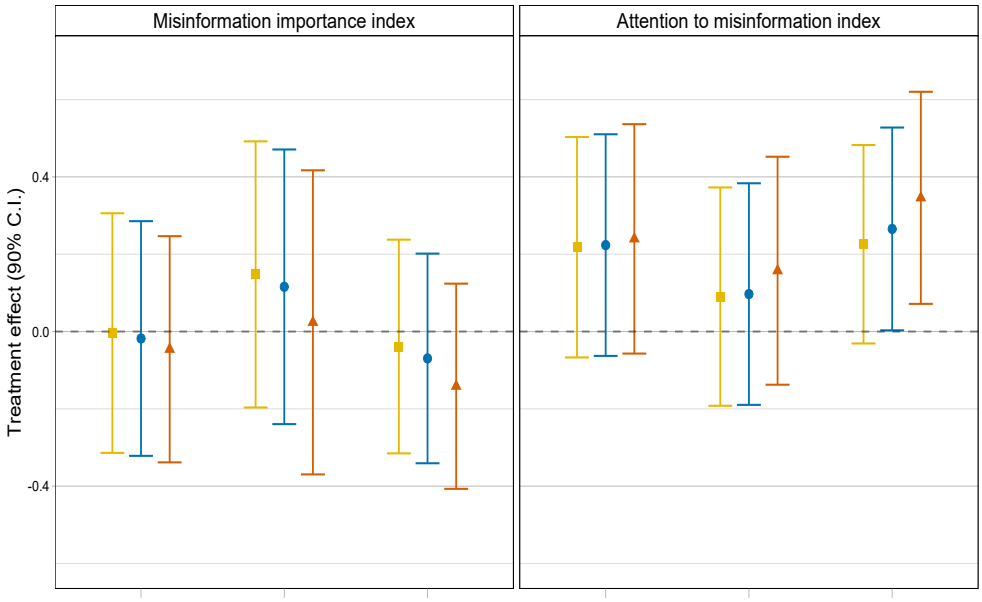
Panel C: Treatment effects with covariates selected with LASSO

Panel C.1: Sample course Delivery Online												
Course and verification	−0.084	0.539***	−0.180	0.099	−0.009	0.209	0.260	0.525***	0.121	−0.203	0.085	0.110
	(0.184)	(0.186)	(0.174)	(0.204)	(0.198)	(0.173)	(0.184)	(0.143)	(0.153)	(0.167)	(0.203)	(0.174)
Course	0.009	0.007	0.313	−0.219	0.065	0.187	0.539***	0.661***	−0.173	−0.162	0.261	0.228
	(0.233)	(0.193)	(0.201)	(0.190)	(0.193)	(0.176)	(0.183)	(0.136)	(0.157)	(0.156)	(0.224)	(0.185)
Verification	−0.149	0.072	−0.179	−0.100	−0.086	0.334**	0.152	−0.035	0.120	−0.260*	0.404**	−0.072
	(0.168)	(0.165)	(0.169)	(0.170)	(0.202)	(0.156)	(0.176)	(0.114)	(0.130)	(0.157)	(0.183)	(0.182)
C.2 Sample Course Delivery WhatsApp												
Course and verification	−0.028	0.085	0.305*	0.173	0.121	0.358*	0.091	0.200	−0.110	0.285**	−0.007	0.169
	(0.147)	(0.133)	(0.160)	(0.142)	(0.156)	(0.193)	(0.148)	(0.123)	(0.154)	(0.143)	(0.137)	(0.151)
Course	−0.226	−0.135	0.139	0.200	0.261*	0.025	0.080	0.100	−0.029	0.414***	0.085	0.025
	(0.155)	(0.138)	(0.163)	(0.148)	(0.153)	(0.211)	(0.129)	(0.121)	(0.180)	(0.133)	(0.147)	(0.155)
Verification	−0.176	0.019	0.215	0.072	0.149	0.046	0.096	0.026	−0.002	0.447***	0.279	0.173
	(0.160)	(0.141)	(0.137)	(0.156)	(0.155)	(0.146)	(0.157)	(0.124)	(0.160)	(0.143)	(0.172)	(0.123)
Outcome mean	0	0	0	0	0	0	0	0	0	0	0	0
Outcome std. dev.	1	1	1	1	1	1	1	1	1	1	1	1
Outcome range	[-3.2,0.6]	[-4.1,3.1]	[-3.8,2.5]	[-2.6,2.2]	[-3.4,1.7]	[-3.6,1.4]	[-3.3,2.5]	[-3.3,2.1]	[-2.8,1.8]	[-1.7,3.3]	[-3,5.1]	[-3,1.3]
Num. LASSO covariates	0	9	8	2	1	0	8	7	12	3	6	7
Observations	463	463	463	463	463	463	463	463	463	463	463	463
R^2	0.480	0.873	0.930	0.799	0.880	0.929	0.636	0.907	0.593	0.646	0.724	0.832

Notes: We report estimates from OLS regression including randomization block fixed effects. All specifications in all panels include the corresponding outcome variables at baseline as a control. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel A and B further include covariates in the outcome family as controls. Specifications in Panel C are from a 10-fold cross validation LASSO model with lambda chosen to be that of the minimum average cross-validated error. The treatment indicators, lagged dependent variable, and fixed effects are forced into the LASSO model and covariates are selected from a pool of 72 baseline variables for each specification. Robust standard errors are in parentheses. * denotes $p<0.1$, ** denotes $p<0.05$, and *** denotes $p<0.01$.

Figure S7: Treatment effect comparison between the OLS, and WLS with restricted Americas Barometer (WLS Americas) and restricted Latin Barometer (WLS Latin) weights on mechanisms: misinformation importance and attention to misinformation

(a) Sample course delivery online



(b) Sample course delivery WhatsApp

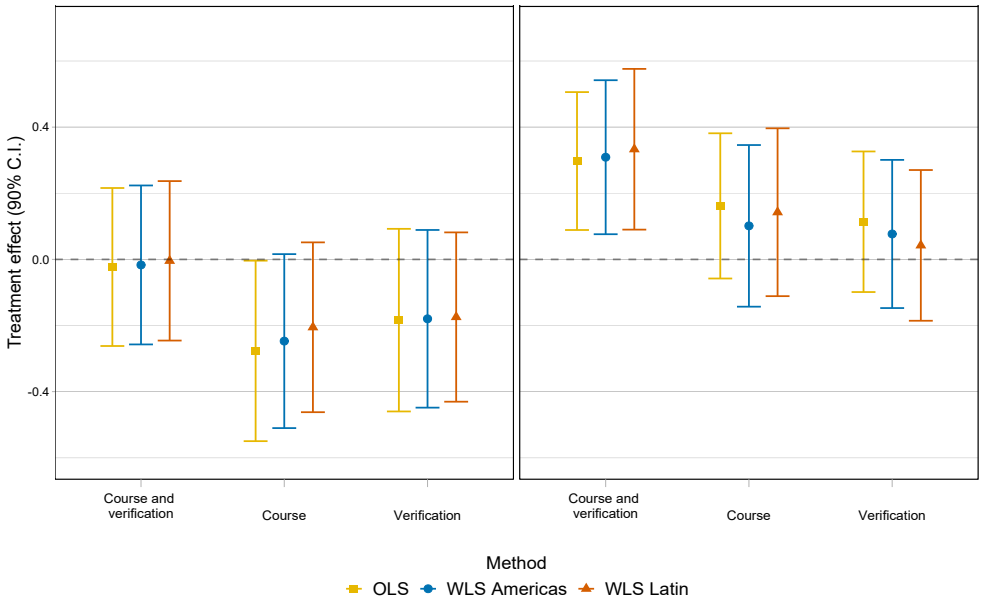
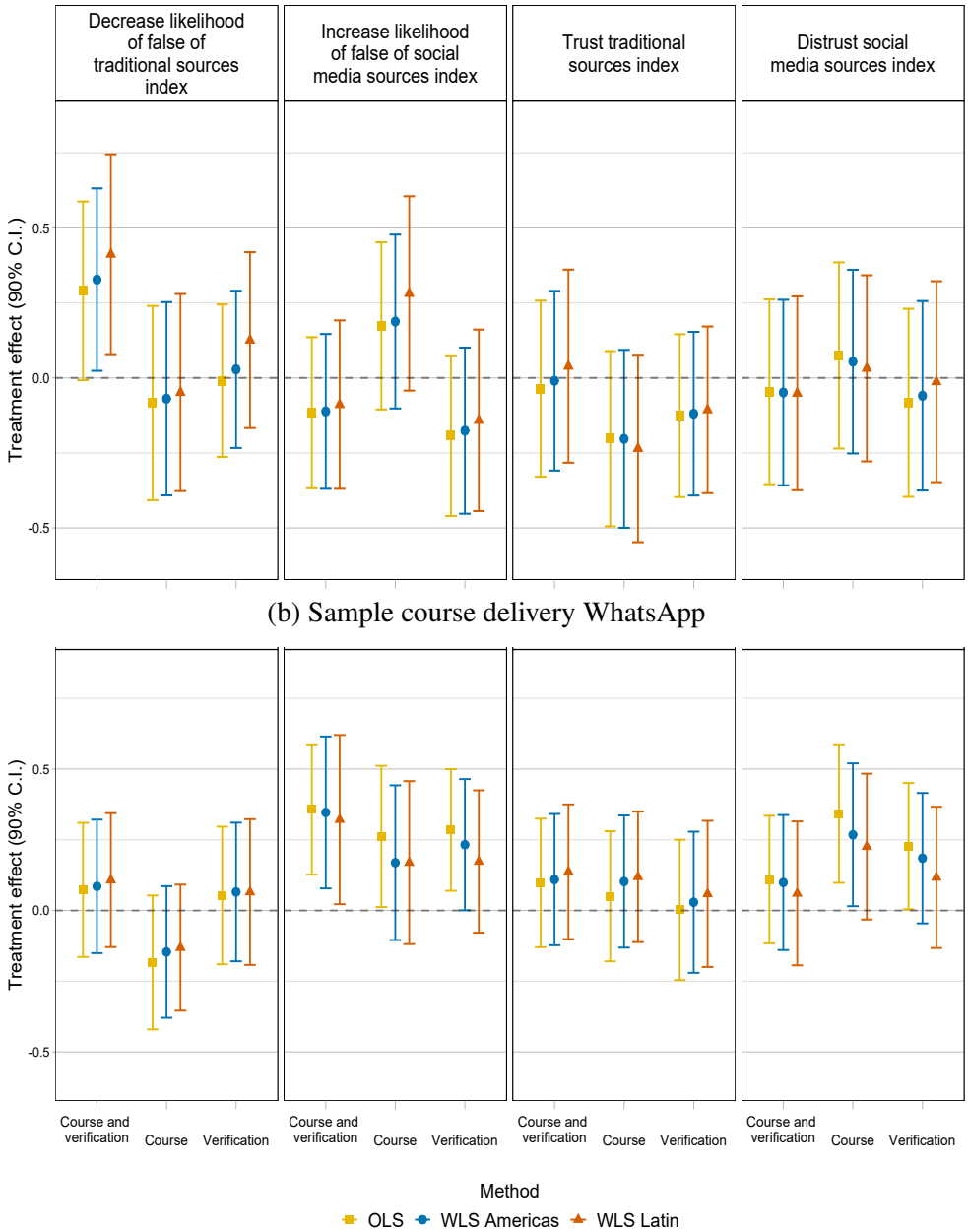


Figure S8: Treatment effect comparison between the OLS, and WLS with restricted Americas Barometer (WLS Americas) and restricted Latin Barometer (WLS Latin) weights on mechanisms: likelihood of false and distrust of traditional and social media sources

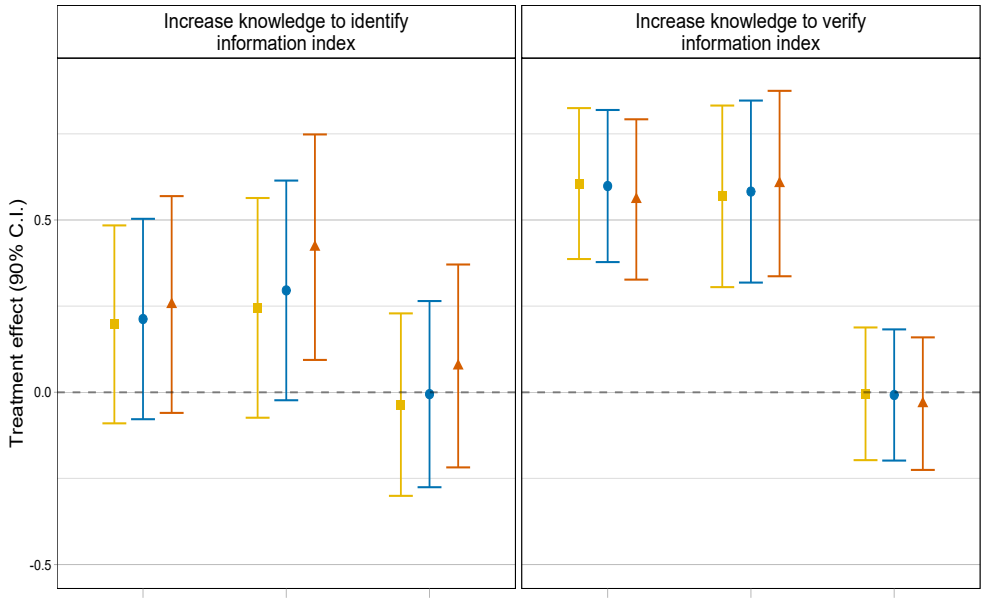
(a) Sample course delivery online



(b) Sample course delivery WhatsApp

Figure S9: Treatment effect comparison between the OLS, and WLS with restricted Americas Barometer (WLS Americas) and restricted Latin Barometer (WLS Latin) weights on knowledge to identify and verify information

(a) Sample course delivery online



(b) Sample course delivery WhatsApp

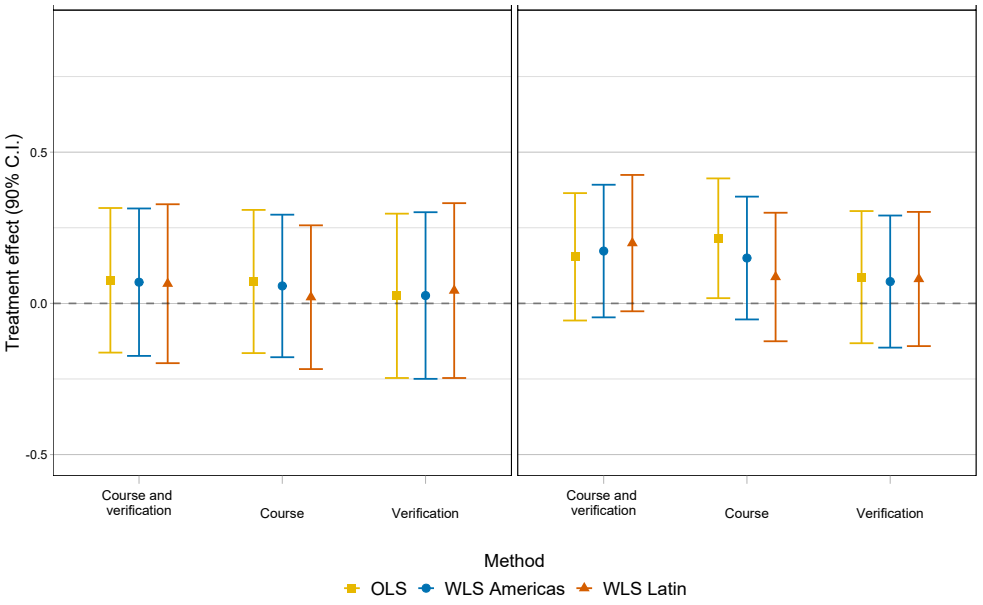
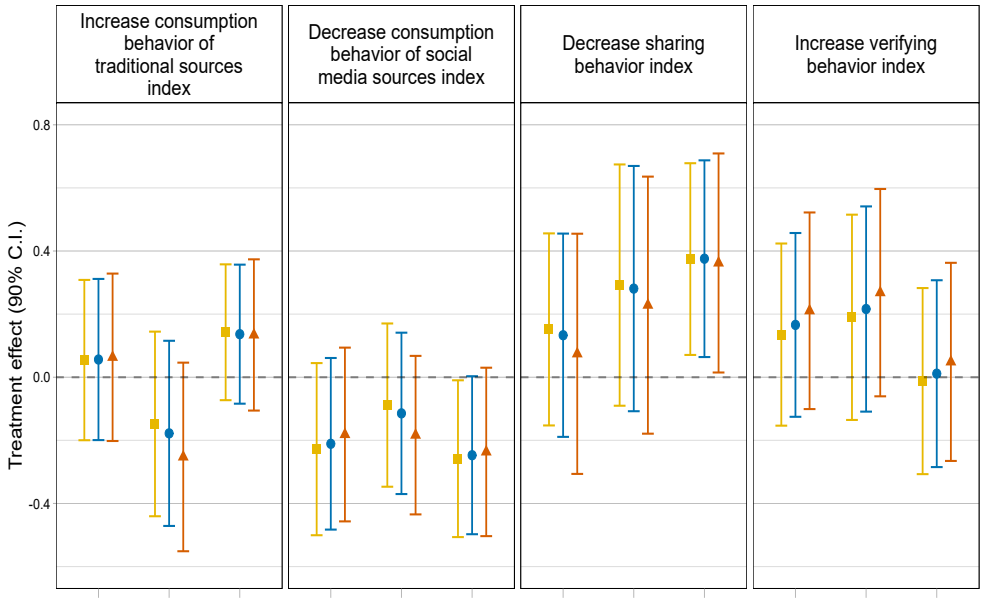
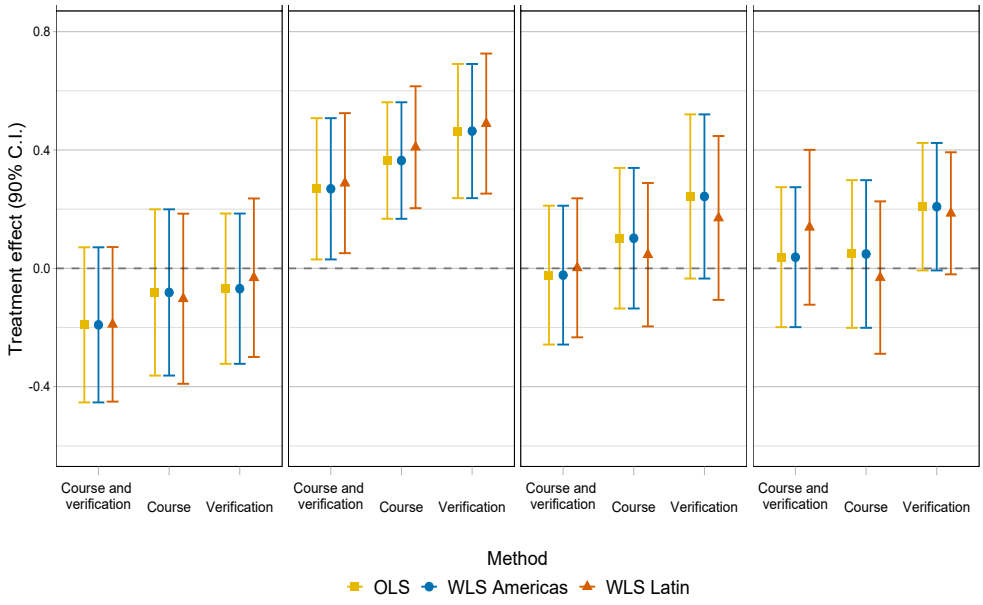


Figure S10: Treatment effect comparison between the OLS, and WLS with restricted Americas Barometer (WLS Americas) and restricted Latin Barometer (WLS Latin) weights on consumption, sharing, and verifying behavior

(a) Sample course delivery online



(b) Sample course delivery WhatsApp



Social media analysis

Table S26: Difference in attrition reporting social media at baseline, posting at least once within one-month after the intervention, and number of posts

	Differential Attrition reporting social media account	Posted at least once within 1 month after the intervention began	Number of posts (log)
	(1)	(2)	(3)
A.1 Sample course Delivery Online			
Course and verification	−0.044	−0.095	−0.183
	(0.074)	(0.102)	(0.353)
Course	0.089	−0.102	−0.373
	(0.077)	(0.103)	(0.325)
Verification	−0.020	−0.074	−0.548
	(0.076)	(0.105)	(0.335)
A.2 Sample Course Delivery WhatsApp			
Course and verification	−0.015	0.082	0.208
	(0.060)	(0.067)	(0.216)
Course	−0.051	−0.039	−0.029
	(0.060)	(0.060)	(0.173)
Verification	−0.078	0.075	0.301
	(0.059)	(0.072)	(0.222)
Outcome mean	0.494	0.321	0.544
Control mean	0.516	0.324	0.566
Observations	838	414	414
R^2	0.095	0.276	0.168

Notes: We report estimates from OLS regression including randomization block fixed effects. Robust standard errors are in parentheses. * denotes p<0.1, ** denotes p<0.05, and *** denotes p<0.01.

Table S27: Treatment Effects on information shared on social media

	Share misinformation index	Share fact-checks
	(1)	(2)
A.1 Sample course Delivery Online		
Course and verification	−0.392*	0.012
	(0.228)	(0.016)
Course	−0.197	0.053
	(0.209)	(0.034)
Verification	−0.385**	−0.018
	(0.188)	(0.016)
A.2 Sample Course Delivery WhatsApp		
Course and verification	0.137	−0.001
	(0.291)	(0.020)
Course	0.037	−0.037
	(0.348)	(0.039)
Verification	0.157	0.041
	(0.371)	(0.030)
Outcome mean	0	0.023
Outcome std. dev.	1	0.149
Outcome range	[-1.1,5.4]	[0,1]
Observations	133	133
R^2	0.716	0.716

Notes: We report estimates from OLS regression including randomization block fixed effects. Specifications further include the corresponding outcome variables at baseline as controls. Robust standard errors are in parentheses. * denotes p<0.1, ** denotes p<0.05, and *** denotes p<0.01.

Table S28: Bayes Factor for statistically insignificant coefficients in our main indexes estimates

	Online sample			WhatsApp sample		
Index	Course and Verification	Course	Verification	Course and Verification	Course	Verification
Misinformation importance	0.313	0.46	0.56	0.106	0.248	0.818
Decrease likelihood of false of traditional sources	0.743	0.434	0.282	0.224	0.119	0.29
Increase likelihood of false of social media sources	0.251	1.515	0.282	-	0.316	-
Trust traditional sources	0.343	0.516	0.292	0.074	0.154	0.305
Distrust social media sources	0.226	0.873	0.288	0.161	-	0.339
Attention to misinformation	2.049	1.072	0.4	-	0.279	0.383
Increase knowledge to identify information	0.337	0.737	0.342	0.088	0.208	0.301
Increase knowledge to verify information	-	-	0.19	584.261	157.431	0.195
Increase consumption behavior of traditional sources	0.329	0.243	0.264	0.121	0.08	0.37
Decrease consumption behavior of social media sources	0.224	0.396	0.221	-	-	-
Decrease sharing behavior	0.422	0.735	-	0.11	0.174	2.067
Increase verifying behavior	0.273	0.246	0.267	0.071	0.088	0.283

Notes: We compute the Bayes Factor for each non-statistically significant coefficient at the 95% level in our main indexes when including all variables in the outcome family (as in panel A of Table S13). The Bayes Factor compares under the null hypothesis the corresponding treatment indicator equal to 0 and under the alternative hypothesis distinct than 0.

Table S29: Power tests for statistically insignificant coefficients in our main indexes estimates

	Online sample			WhatsApp sample		
Index	Course and Verification	Course	Verification	Course and Verification	Course	Verification
Misinformation importance	0.546	0.592	0.506	0.493	0.491	0.516
Decrease likelihood of false of traditional sources	0.546	0.592	0.506	0.493	0.491	0.516
Increase likelihood of false of social media sources	0.546	0.592	0.506	-	0.491	-
Trust traditional sources	0.546	0.592	0.506	0.493	0.491	0.516
Distrust social media sources	0.546	0.592	0.506	0.493	-	0.516
Attention to misinformation	0.546	0.592	0.506	-	0.491	0.516
Increase knowledge to identify information	0.546	0.592	0.506	0.493	0.491	0.516
Increase knowledge to verify information	-	-	0.506	0.493	0.491	0.516
Increase consumption behavior of traditional sources	0.546	0.592	0.506	0.493	0.491	0.516
Decrease consumption behavior of social media sources	0.546	0.592	0.506	0.493	-	-
Decrease sharing behavior	0.546	0.592	-	0.493	0.491	0.516
Increase verifying behavior	0.546	0.592	0.506	0.493	0.491	0.516

Notes: We present the minimum detectable effect given our sample size, a significance level of 0.05, and power of 0.80. Throughout, we perform a two-sided test.

Table S30: Power tests for statistically insignificant coefficients in our main indexes estimates

	Online sample			WhatsApp sample		
Index	Course and Verification	Course	Verification	Course and Verification	Course	Verification
Misinformation importance	0.484	0.523	0.448	0.437	0.435	0.457
Decrease likelihood of false of traditional sources	0.484	0.523	0.448	0.437	0.435	0.457
Increase likelihood of false of social media sources	0.484	0.523	0.448	-	0.435	-
Trust traditional sources	0.484	0.523	0.448	0.437	0.435	0.457
Distrust social media sources	0.484	0.523	0.448	0.437	-	0.457
Attention to misinformation	0.484	0.523	0.448	-	0.435	0.457
Increase knowledge to identify information	0.484	0.523	0.448	0.437	0.435	0.457
Increase knowledge to verify information	-	-	0.448	0.437	0.435	0.457
Increase consumption behavior of traditional sources	0.484	0.523	0.448	0.437	0.435	0.457
Decrease consumption behavior of social media sources	0.484	0.523	0.448	0.437	-	-
Decrease sharing behavior	0.484	0.523	-	0.437	0.435	0.457
Increase verifying behavior	0.484	0.523	0.448	0.437	0.435	0.457

Notes: We present the minimum detectable effect given our sample size, a significance level of 0.10, and power of 0.80. Throughout, we perform a two-sided test.