HOW EFFECTIVE ARE MEDIA LITERACY INTERVENTIONS AT COUNTERING MISINFORMATION IN THE GLOBAL SOUTH?

Robert A. Blair, Jessica Gottlieb, Brendan Nyhan, Laura Paler, Charlene J. Stainfield, and Julie Anne Weaver

Media literacy interventions are an increasingly popular approach to countering the spread of misinformation. Yet, until recently, the evidence base for assessing their effectiveness in Global South countries has been limited.[1] This brief, based on a longer report,[2] identifies, reviews, and synthesizes the results from nine recent randomized control trials of media literacy interventions conducted in the Global South to produce actionable policy recommendations.

BACKGROUND

MEDIA LITERACY INTERVENTIONS ARE CENTRAL TO THE FIGHT AGAINST MISINFORMATION

Misinformation presents a serious threat to democracy, social cohesion, trust in institutions, and public health worldwide.[3] While the spread of misinformation has received substantial attention in Global North countries, the danger it poses to Global South countries is just as acute. For example, misinformation on WhatsApp is implicated in the widespread questioning of electoral integrity and the media in Brazil;[4] the

Highlights

Media literacy interventions are effective in the Global South in some, but not all, cases. Because results are mixed, it is especially imperative to understand when and under what conditions they are most effective and where additional research is needed to know how best to allocate resources.

Media literacy interventions as implemented and tested in the Global South appear to be more effective among individuals with higher levels of education and digital literacy. Policymakers and practitioners targeting individuals with lower digital literacy should consider evaluating how additional measures or approaches might be more effective with these populations.

Simply teaching new media literacy techniques may be insufficient to change attitudes and behaviors, especially when the misinformation relates to group attachments or deeply held beliefs. Some studies suggest media literacy interventions are more effective when they not only teach skills to counteract misinformation but also motivate individuals to use those skills by appealing to emotions, social norms, or group attachments.

Intensive media literacy interventions are not necessarily more effective than light-touch interventions at improving discernment between true and false information or intent to share misinformation. And **neither light-touch nor intensive interventions reliably generate durable improvements in outcomes**. Policymakers and practitioners should, therefore, consider deploying light-touch interventions that can scale more cost-effectively immediately prior to important, misinformation-prone events (e.g., elections).

Box 1: Defining Media Literacy Interventions to Counter Misinformation

Media literacy interventions provide individuals with broad tools and skills to identify and resist misinformation. This focus on general skills differentiates media literacy from other information-based interventions to combat misinformation, like inoculation and debunking, which seek to correct specific false or misleading claims. Media literacy interventions thus mostly rest on the assumption that individuals believe and share misinformation because they lack the ability to recognize it as false. Some media literacy interventions also consider how factors such as emotions or group attachments affect peoples' ability and willingness to resist misinformation.

Though they share a focus on building skills, the media literacy interventions tested in the studies we review vary substantially in their form. *Light-touch* initiatives are one-time brief exposures, such as presenting a tip sheet or infographic to be consumed by individuals at their own pace. *Intermediate* interventions go a step beyond this, either by repeating short lessons over time or by adding additional elements, like videos or feedback. *Intensive* interventions involve more extensive in-person training by research staff or experts and/or take place in multiple sessions over days or weeks.

amplification of hate speech and misinformation on Facebook is believed to have spurred ethnic violence targeting minority groups in Ethiopia and Myanmar;[5, 6] and a falsely attributed video is thought to have contributed to ethnic violence in Côte d'Ivoire.[7]

Media literacy interventions are increasingly prominent among efforts to curb the spread of misinformation in the Global South. In 2021, the United Nations General Assembly adopted a resolution that called for nations "to develop and implement policies, action

plans and strategies related to the promotion of media and information literacy, and to increase awareness, capacity for prevention and resilience to disinformation and misinformation."[8] In response to this call, UNESCO deployed 26 distinct media literacy programs in 59 countries/regions, including 54 in the Global South, with a total budget of just under five million USD between 2022 and 2023.[9] Media literacy interventions have been supported and implemented by a number of other governmental and non-governmental organizations, including the United States Agency for International Development (USAID), the National Institutes of Health (NIH), IREX, and the International Development Research Centre (IDRC).

EXPERTS VIEW MEDIA LITERACY INTERVENTIONS AS ESPECIALLY PROMISING

Experts perceive media literacy interventions to be a promising approach to combating misinformation in the Global South. We conducted a survey of 138 policymakers, practitioners and academic experts, and asked respondents to select the initiatives that they expected would be most effective at countering misinformation in the Global South. Experts on the Global South viewed media literacy as the most promising approach; experts on misinformation in the Global North saw it as the second most promising approach (see Figure 1).[10]

THERE IS AN OPTIMISM-EVIDENCE GAP ON THE EFFECTIVENESS OF MEDIA LITERACY INTERVENTIONS

Though media literacy interventions have become increasingly popular, relatively few rigorous studies have evaluated their

Figure 1: Expert assessments of the most effective intervention types to counter misinformation in the Global South

Allocations across 12 intervention types

Mean allocation
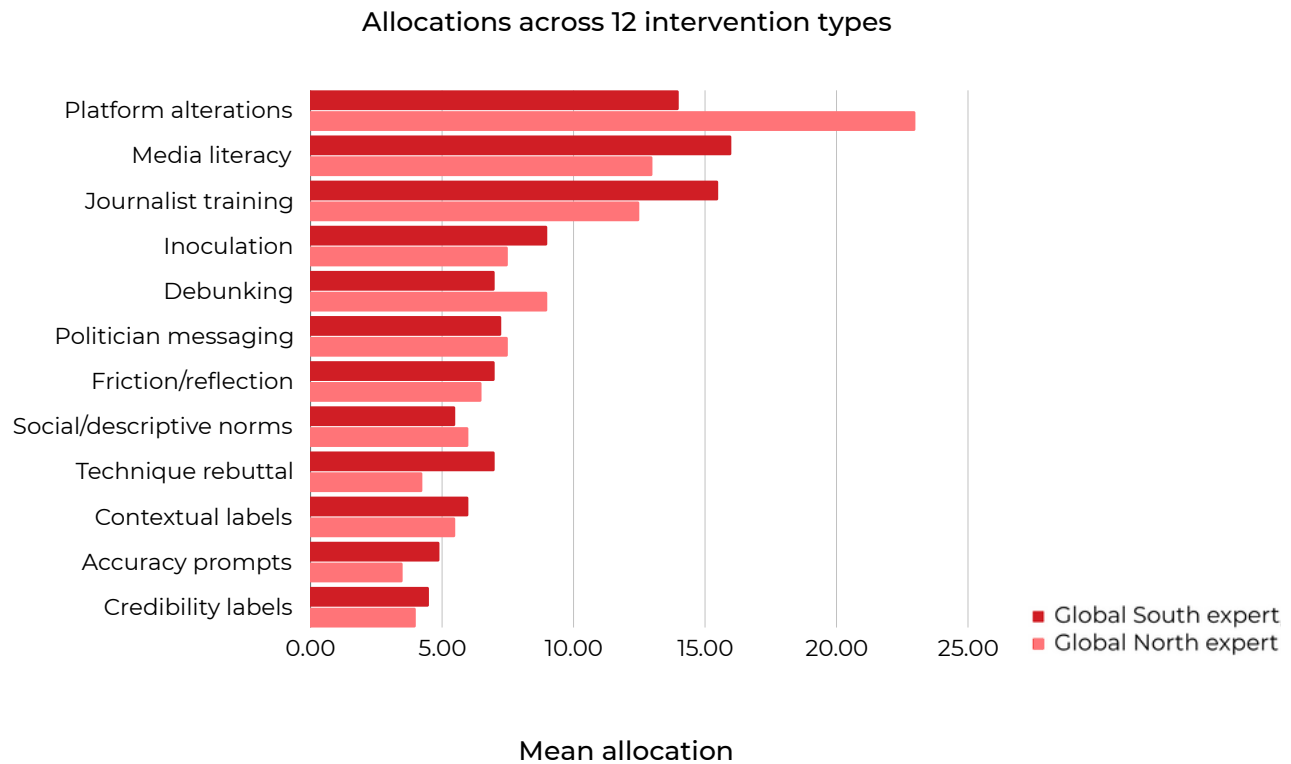
Source: Results from an original survey of 138 practitioners, policymakers, and academic researchers (conducted in April 2023).[11] All respondents were asked how they would allocate 100 units of funding across the 12 intervention types depicted in the graph to indicate which they thought would be most effective at countering misinformation in the Global South.

effectiveness in a way that allows us to attribute causal effects. According to a recent report,[12] only 16 randomized control trials have been conducted on media literacy interventions. There is much more evidence on the effectiveness of other informational interventions to counter misinformation, including inoculation/pre-bunking (25 studies), debunking (56 studies), and credibility labels (24 studies).

Evidence on the effectiveness of media literacy interventions for the Global South is even

thinner when considering that, of the 16 total studies identified in that report, nine focused on Global North countries. While evidence from Global North countries can be informative for understanding whether media literacy interventions might work in the Global South, it cannot substitute for direct evidence from Global South countries. Fortunately, a number of recent high-quality studies focused on the Global South now provide a foundation for assessing the evidence not only on whether media literacy interventions work but also on when and why they are most effective.

Box 2: Why randomized control trials?

We limit our review to randomized control trials (RCTs) because of the fundamentally causal nature of the question this brief addresses. This brief aims to assess whether media literacy interventions cause changes in outcomes like intent to share misinformation. Identifying a causal relationship requires isolating the effects of a media literacy intervention from all other factors that might affect the outcomes of interest.

Isolating an intervention's effect can be extremely challenging in studies that do not randomly assign participants to treatment and control groups. Consider two scenarios:

- An organization implements a media literacy intervention and allows individuals (or communities) to volunteer to participate. Afterward, the organization compares outcomes among those who wanted to participate to a similar group that did not. However, those who want to participate in a media literacy intervention likely differ from those who do not on many factors (e.g., education levels, income, motivation to combat misinformation), making it impossible to know for sure whether any differences in outcomes are due to the intervention or to these other factors.

- An organization implements a multi-day media literacy intervention and, for all participants, measures discernment before, immediately after, and

several weeks after the intervention. The organization then compares the post-intervention outcomes to the pre-intervention measure. Unfortunately, this approach cannot rule out the possibility that other events or experiences during the study period could also have affected participant outcomes.

Because RCTs produce a control group that resembles the treatment group, they have high *internal validity,* meaning they do a good job estimating causal effects in their immediate study context. One concern about RCTs is that they lack *external validity*, meaning results might not generalize to other contexts or populations. This challenge is not unique to RCTs; it applies when trying to extrapolate results from any study—qualitative or quantitative—that draws evidence from specific contexts or populations. See Box 5 for more on how to think about extrapolating RCT evidence on media literacy interventions to different settings and populations.

Importantly, other forms of evidence, such as non-randomized impact evaluations or qualitative case studies, can also provide valuable insights into the effects of media literacy interventions. We do not include these studies in the evidence review because they are more limited than RCTs in their ability to isolate the causal effects of interventions. They can, however, provide valuable insights into mechanisms, help to interpret findings, and generate new ideas to study.

METHODOLOGY

We conducted a comprehensive search of all randomized control trials (RCTs) of media literacy interventions in the Global South (see Box 2 on RCTs).[13] The search process is outlined in greater detail in the Supplemental Information section. We restrict our search to studies that evaluate media literacy interventions on two outcomes: (1) discernment, which is the ability to distinguish true from false information; and/or (2) the intent to share false (relative to true) information. This is important for figuring out

not only whether media literacy interventions reduce engagement with false information but also whether they have an unintended harmful effect of making people more skeptical of truthful information as well.

The search identified nine studies testing nine interventions spanning 14 Global South countries.[14] The studies provide relatively good geographic coverage, including South Asia; East Asia and the Pacific; Sub-Saharan Africa; Latin America and the Caribbean; Europe and Central Asia; and the Middle East and North Africa. Boxes 3 and 4 offer additional details on two of the studies, showing an

example of both a more intensive and a more light-touch initiative and highlighting a case of finding positive versus no effects. A list of all articles can be found in Table 1 in the Supplemental Information, with additional details available in an online [database](#).[15]

<div style="background:#e8404c;color:white;padding:1em">

Box 3: Evaluating in-person media literacy trainings in India[16]

</div>

This study tests the effects of a more intensive in-person training in India and highlights some of the limitations in the effectiveness of media literacy interventions. The training involved a one-hour session conducted one-on-one in participants' homes, where participants went through a learning module on how to verify information, practiced fact-checking four false stories, and learned tips to reduce the spread of misinformation.

Researchers conducted the study with a random sample of 1,224 individuals in the city of Gaya in Bihar, a state with one of the lowest literacy rates in India. The intervention was implemented in the lead-up to the 2019 elections when politics was highly salient and misinformation was pervasive.

Participants were randomly assigned to either a control group (that received a placebo intervention) or one of two media literacy treatment groups (the groups only differed in whether the fact-checking exercise featured information that was pro- or anti-BJP, the leading right-leaning national party in India). The researcher measured the effects of the intervention on discernment two weeks after the training took place.

So was it effective? The study finds that the media literacy training produced no effect on average, meaning that those who participated in the training were no more or less able to classify true or false information correctly than those in the control group. However, intervention effects vary for individuals who are and are not supporters of the BJP. Specifically, the initiative increased the ability of non-BJP supporters to identify false information but reduced the ability of BJP supporters to do the same. These surprising adverse effects were most pronounced for pro-BJP supporters when asked to identify pro-BJP stories as false. Overall, this study shows that individuals can be highly resistant to efforts to counter pro-attitudinal misinformation, and provides support for a growing line of research suggesting that psychological factors like group attachments are a key barrier to interventions' effectiveness.

MAIN TAKEAWAYS

TAKEAWAY 1:
Media literacy interventions work in the Global South but not all of the time.

Findings are mixed across the interventions evaluated. Overall, the studies find positive effects for four of the nine media literacy interventions evaluated.[17] Two of the media literacy interventions produced no effects,[18] while the remaining three interventions yielded mixed results, meaning some combination of positive, negative, or null findings.[19] Of the four media literacy interventions that had positive effects, one was conducted with media and communication college students in Nigeria,[20] raising questions about whether results might generalize to broader populations. Importantly, none of the interventions caused adverse effects (e.g., a reduced ability to identify false information or increased willingness to share false information) on average, although one study reports adverse effects on a specific group (see Box 3). Because the overall results are mixed, it is especially important to identify the intervention designs and populations for which media literacy programs are most effective, while recognizing that some interventions have not been thoroughly evaluated.

TAKEAWAY 2:
Media literacy interventions tend to be more effective for individuals with more education and digital literacy.

While some studies show that media literacy interventions work in general, findings from three of the nine interventions suggest that media literacy programs might have larger effects on individuals with higher levels of

Box 4: Comparing text-based media literacy interventions in Kenya[21]

In this less intensive intervention from Kenya, researchers compared different approaches to see which worked best. The researchers collaborated with a local NGO to create text message media literacy 'courses' with messages that were sent to participants once a day for five days. Participants were recruited from ads placed on Facebook and received a small monetary incentive to participate. Around 9,000 people participated in the study.

Researchers randomly assigned participants to either a control group or one of three treatment groups featuring a different text message course. One course taught participants how to identify misleading reasoning in misinformation while another helped people identify how misinformation manipulates emotions. The third course combined both the reasoning and emotions treatments. This set-up allows the researchers to test not only the effectiveness of any one course but also to compare the effectiveness of the different courses.

Researchers measured effects on both discernment and sharing intentions using online surveys. Outcomes were measured both immediately after the intervention and 7-11 weeks later, allowing them to assess the durability of effects.

What did the researchers find? All three text message courses improved discernment and reduced participants' intent to share misinformation. All three treatments were effective both immediately following the intervention and 7-11 weeks later, suggesting that the effects are durable. While all three text message courses improved outcomes, the course that emphasized how misinformation manipulates emotions had a larger effect size than the reasoning course and the one that combined the two approaches.

education and digital literacy.[22] Only one of these studies directly examines whether the effects of its media literacy intervention vary across individuals, finding that the positive effects of the intervention were driven almost entirely by those with high digital literacy.[23] Moreover, an online intervention conducted with a more educated and digitally literate population in India was found to be effective,[24] whereas two in-person interventions conducted with less educated, rural populations in India had no effect.[25] It is also worth noting that one intervention that produced positive effects was conducted only on samples of college students in Nigeria.[26] Because they only include participants with high education, the Nigeria studies cannot speak to whether media literacy interventions are more effective for individuals with higher versus lower levels of education. However, the results do tell us, at a minimum, that media literacy interventions work with highly educated populations.

These findings raise important questions for policymakers and practitioners about the target populations they are seeking to influence with media literacy interventions. If the goal is to reach less educated people with lower levels of digital literacy, the evidence suggests current approaches that have been tested here are not working, and that other strategies need to be rigorously evaluated.

TAKEAWAY 3:
Media literacy interventions appear to be more effective when they not only provide individuals with the skills to counteract misinformation but also strengthen the will to use those skills.

The majority of studies evaluate standard media literacy interventions, which aim to equip participants with the skills to distinguish true from false information and to reduce their sharing of misinformation. These studies,

however, often do not consider how other factors—such as emotions, political incentives, and social pressure—can act as barriers to acquiring and acting on these skills.

Three of the nine studies provide important exceptions by examining how media literacy programs could be more effective if they address non-skills barriers to countering misinformation. These studies show how addressing emotions,[27] political incentives to share misinformation,[28] and ethnic polarization can impact the effectiveness of media literacy interventions.[29] Taken together, these studies provide noteworthy evidence that combining standard media literacy interventions with approaches that target individuals' motivation to use their new skills could be a productive strategy.

TAKEAWAY 4:
In the Global South, there is no evidence that intensive media literacy interventions are more effective than light-touch interventions.

The designs of the media literacy interventions reviewed here vary substantially. Five of the nine interventions are light-touch.[30] Two of the nine interventions are more intermediate, such as a daily text message sent over five days[31] or feedback from research staff.[32] Finally, two of the nine interventions are intensive in that they were implemented in-person and for a longer duration.[33]

There is no evidence that more intensive interventions are more effective. Of the four interventions with positive effects, two were light-touch,[34] one was intermediate,[35] and one was intensive.[36] In one exception, researchers found that a light-touch version of the intervention, which involved watching a short video, was not effective unless it was coupled with a more intensive version of the treatment that involved personalized feedback on

discernment performance.[37] Still, on the whole, the evidence does not support the conclusion that more intensive interventions are more effective.

At the same time, the interventions reviewed in this brief are not as intensive as some media literacy programming carried out around the world, which can take the form of training that lasts weeks at a time or curriculum embedded in broader civic or adult education courses. More rigorous evaluations of such media literacy interventions would help identify whether such higher-cost, very intensive interventions are worthy investments compared to the kind of interventions studied here.

TAKEAWAY 5:
There is little evidence that media literacy interventions in the Global South, whether light touch or more intensive, produce longer term or durable outcomes.

Of the nine interventions, three are evaluated using only immediate outcomes, meaning outcomes are measured immediately after the intervention concludes;[38] two are assessed only with longer term outcomes sometime after the intervention;[39] and four are studied by measuring both.[40]

Only studies that measure both immediate and longer term outcomes can address the question of durability, which requires identifying whether an intervention generated an initial effect that endured over time. Of the four interventions that were evaluated in terms of both immediate and longer term outcomes, only one produced an initial positive effect that persisted over time.[41] Two interventions generated initially positive effects that did not endure,[42] while the second intervention tested in one study generated neither an immediate nor a longer term effect.

In addition, of the six interventions that were studied by measuring only longer term outcomes, only one produced positive effects detected seven to eleven weeks after the intervention.[43] The remaining five interventions yielded no longer term effects.

Overall, the studies that measure longer term outcomes include both light-touch and intensive interventions. We find no evidence to support the expectation that intensive interventions produce longer term or more durable outcomes. However, even the more intensive interventions studied here were still less intensive than some programs being currently implemented.

RECOMMENDATIONS

This evidence review suggests the following three actionable policy recommendations for the design of media literacy interventions going forward:

1. **Consider the target population of interest when designing media literacy interventions**. At present, the evidence suggests that media literacy interventions might work best on populations that are already relatively well-educated and have high levels of digital literacy. One option would be to prioritize these populations where media literacy interventions are likely to be more effective. If the goal is to instead reach those with lower levels of education or digital literacy, additional measures need to be evaluated.

2. **Consider factoring the psychology of misinformation and the role of emotions into the design of media literacy interventions.** The evidence reviewed here suggests that emotions and the drive to reinforce pre-existing beliefs can play an

---

Box 5: Do findings generalize?

A common challenge for RCTs (as well as for other types of research) is determining whether evidence from particular studies generalizes to other populations, contexts, or intervention designs. For instance, if a study shows that a media literacy intervention worked on an educated population in India, would the same intervention produce similar results on educated populations in Southeast Asia or Sub-Saharan Africa? Alternatively, would a slightly different intervention produce similar results on the same educated population in India?

These kinds of questions are important to ask but hard to answer definitively. Nevertheless, the following questions can guide efforts to think through when, why, and how results from one setting might be informative for other settings.

1. Are there studies conducted in places that have relevant **country-level characteristics** that are similar to the setting of interest? If not, is there any reason to believe that differences in country-level characteristics could produce different outcomes?

2. Are there studies conducted on populations that have similar **population-level characteristics** to the population of interest? If not, is there good reason to believe that your population of interest differs in ways that could produce different outcomes?

3. Are there studies that evaluate **similar interventions** to the intervention of interest? If not, is there any reason why your intervention design differs in ways that would produce different outcomes?

Evidence collected in settings with similar country characteristics, population characteristics, and intervention designs is best but does not always exist. To the extent that prior research differs on one or more of these dimensions, the goal should be to consider whether and how these differences might affect the expected results.

---

important role in individuals' motivation to use the skills taught by standard media literacy interventions. Interventions that also address the underlying motivation to discern true from false information or to share information people suspect is false might be more effective than interventions that focus on imparting skills alone.

3. **Design and implement interventions knowing that effects will likely be short-lived**. A consistent finding across the studies reviewed here is that even relatively more intensive interventions generally did not produce durable results. Consequently, policymakers and practitioners should (1) strategically time media literacy interventions to take place immediately prior to key events (e.g., elections); (2) prioritize light-touch interventions that could be repeated over time and scaled more easily; or (3) combine traditional media literacy interventions with other elements (e.g., emotional appeals) to enhance the durability of effects.

ADDITIONAL RESOURCES

Access the full USAID report on Misinformation, on which this evidence brief is based: https://pdf.usaid.gov/pdf_docs/PA0215JW.pdf

Search the database of studies consulted in the full USAID report on Misinformation: https://www.democratic-erosion.com/briefs/misinformation-intervention-database/

Find out more about each of the studies included in this evidence brief: https://www.democratic-erosion.com/media-literacy-database/

ABOUT THE DEMOCRATIC EROSION CONSORTIUM

The Democratic Erosion Consortium is a partnership of researchers, students, policymakers, and practitioners committed to marshaling evidence and learning to address the growing crisis of democratic erosion worldwide.

Learn about our other Democratic Erosion Evidence Briefs: https://www.democratic-erosion.org/briefs/

Stay up to date by joining the Democratic Erosion Consortium list-serve: https://www.democratic-erosion.org/join-our-listserv/

About the Authors

Author Affiliations

Robert A. Blair - Arkadij Eisler Goldman Sachs Associate Professor, Department of Political Science and Watson Institute for International and Public Affairs, Brown University
Jessica Gottlieb - Associate Professor, Hobby School of Public Affairs, University of Houston
Brendan Nyhan - James O. Freedman Presidential Professor, Department of Government, Dartmouth College
Laura Paler - Associate Professor, Department of Government, School of Public Affairs, American University
Charlene J. Stainfield - Ph.D. Candidate, Department of Political Science, Ohio State University
Julie Anne Weaver - Research Director, Democratic Erosion Consortium

ENDNOTES

[1] By Global South, we mean the low- and middle-income countries, following the World Bank's income designations, in East Asia and the Pacific, Eastern Europe, Central Asia, Latin America and the Caribbean, the Middle East and North Africa, South Asia, and Sub- Saharan Africa.

[2] This brief is based on a longer report commissioned by USAID synthesizing evidence on strategies to combat misinformation in both the Global North and Global South. Blair, Robert A., Jessica Gottlieb, Brendan Nyhan, Laura Paler, Pablo Argote, and Charlene J. Stainfield. "Interventions to Counter Misinformation: Lessons from the Global North and Applications to the Global South." USAID (2023). https://pdf.usaid.gov/pdf_docs/PA0215JW.pdf. The report is also summarized in: Blair, Robert A., Gottlieb, Jessica, Nyhan, Brendan, Paler, Laura, Argote, Pablo, and Charlene J. Stainfield. 2024. "Interventions to Counter Misinformation: Lessons from the Global North and applications to the Global South." Current Opinion in Psychology 55:101732. https://doi.org/10.1016/j.copsyc.2023.101732.

[3] Lewandowsky, Stephan, Ecker, Ullrich K.H., Cook, John, van der Linden, Sander, Roozenbeek, Jon, and Naomi Oreskes. 2023. "Misinformation and the epistemic integrity of democracy." Current Opinion in Psychology 54:101711. https://doi.org/10.1016/j.copsyc.2023.101711.

[4] Bandeira, L., D. Barojan, R. Braga, J. L. Peñarredonda, and M. F. P. Argüello (2019). Disinformation in democracies: Strengthening digital resilience in Latin America. Atlantic Council, March 28, 2019. Downloaded June 27, 2023 from https://www.atlanticcouncil.org/in-depth-research-reports/report/disinformation-democracies-strengthening- digital-resilience-latin-america.

[5] Jackson, J., L. Kassa, and M. Townsend (2022). Facebook 'lets vigilantes in Ethiopia incite ethnic killing'. February 20, 2022. Downloaded June 27, 2023 from https://www.theguardian.com/technology/2022/feb/20/ facebook-lets-vigilantes-in-ethiopia-incite-ethnic-killing.

[6] Amnesty International (2022). The social atrocity: Meta and the right to remedy for the Rohingya. September 29, 2022. Downloaded June 27, 2023 from https://www.amnesty.org/en/documents/ASA16/5933/2022/en/.

[7] APA News (2021). I/coast: Skirmishes ignited by fake news leave 10 wounded. May 20, 2021. Downloaded June 27, 2023 from https://apanews.net/2021/05/20/icoast-ethnic-clashes-caused- by-fake-news-leave-10-people-wounded/.

[8] UNGA Resolution: Global Media and Information Literacy Week. 2021. A/RES/75/267

[9] UNESCO, Core Data Portal. Accessed December 2023.

[10] Blair, et al., "Interventions to Counter Misinformation."

[11] Ibid.

[12] Ibid.

[13] This updated the search process specifically for media literacy interventions in the Global South that had been conducted for the broader misinformation evidence review we carried out in Blair, et al., "Interventions to Counter Misinformation."

[14] One of the studies (Guess et al 2020) tests two different interventions (one online and one in-person), while two different studies test the same intervention (Apuke, Omar, and Tunca 2022 and Zhang et al. 2022). Guess, Andrew M., et al. "A digital media literacy intervention increases discernment between mainstream and false news in the United States and India." Proceedings of the National Academy of Sciences 117, no. 27 (2020). https://doi.org/10.1073/pnas.1920498117. Apuke, Oberiri Destiny, Bahiyah Omar, and Elif Asude Tunca. "Literacy concepts as an intervention strategy for improving fake news knowledge, detection skills, and curtailing the tendency to share fake news in Nigeria." Child & Youth Services 44, no. 1 (2023). https://doi.org/10.1080/0145935X.2021.2024758 and Zhang, Leilei, et al. "Experimenting the effect of using visual multimedia intervention to inculcate social media literacy skills to tackle fake news." Journal of Information Science (2022). https://doi.org/10.1177/01655515221131797.

[15] https://www.democratic-erosion.com/media-literacy-database/.

[16] Badrinathan, Sumitra. "Educative Interventions to Combat Misinformation: Evidence from a Field Experiment in India." American Political Science Review 115, no. 4 (2021). doi:10.1017/S0003055421000459.

[17] One of the interventions in Guess, et al., "A Digital Media Literacy Intervention" has positive effects. Apuke, Omar, and Tunca, "Literacy Concepts," and Zhang, et al., "Experimenting the Effect" test the same intervention. Athey, Susan Carleton, Matias Cersosimo, Kristine Koutout, and Zelin Li. "Emotion- versus Reasoning-Based Drivers of Misinformation Sharing: A Field Experiment Using Text Message Courses in Kenya." Stanford University Graduate School of Business Research Paper No. 4489759 (2023). http://dx.doi.org/10.2139/ssrn.4489759;

[17 cont.] Apuke, Omar, and Tunca, "Literacy Concepts as an Intervention Strategy," and Zhang, et al., "Experimenting the Effect"; Guess, et al., "A Digital Media Literacy Intervention"; Offer-Westort, Molly, Leah R. Rosenzweig, and Susan Athey. "Battling the Coronavirus Infodemic Among Social Media Users in Africa." arXiv preprint arXiv:2212.13638 (2022). https://doi.org/10.48550/arXiv.2212.13638.

[18] This includes the intervention in Gottlieb, Jessica, Claire Adida, and Richard Moussa. "Reducing Misinformation in a Polarized Context: Experimental Evidence from Côte D'ivoire." OSF Preprints (2023). https://doi.org/10.31219/osf.io/6x4wy, and one of the two interventions tested in Guess, et al., "A Digital Media Literacy Intervention."

[19] Ali, Ayesha, and Ihsan Ayyub Qazi. "Countering misinformation on social media through educational interventions: Evidence from a randomized experiment in Pakistan." Journal of Development Economics 163 (2023): 103108. https://www.sciencedirect.com/science/article/abs/pii/S0304387823000639?via%3Dihub; Arechar, Antonio A., et al. "Understanding and combatting misinformation across 16 countries on six continents." Nature Human Behaviour 7, no. 9 (2023). https://www.nature.com/articles/s41562-023-01641-6; Badrinathan, "Educative Interventions."

[20] Apuke, Omar, and Tunca, "Literacy Concepts," and Zhang, et al., "Experimenting the Effect of Using Visual Multimedia Intervention."

[21] Athey, et al., "Emotion- versus Reasoning-Based Drivers."

[22] Ali and Qazi, "Countering Misinformation" ; Badrinathan, "Educative Interventions" ; Guess, et al., "A Digital Media Literacy Intervention."

[23] Ali and Qazi, "Countering Misinformation."

[24] Guess, et al., "A Digital Media Literacy Intervention."

[25] Badrinathan, "Educative Interventions"; Guess, et al., "A Digital Media Literacy Intervention"

[26] Apuke, Omar, and Tunca, "Literacy Concepts as an Intervention Strategy," and Zhang, et al., "Experimenting the Effect."

[27] Athey, et al., "Emotion- versus Reasoning-Based Drivers."

[28] Badrinathan, "Educative Interventions."

[29] Gottlieb, Jessica, Claire Adida, and Richard Moussa. "Reducing Misinformation."

[30] Arechar, et al., "Understanding and Combatting Misinformation"; Gottlieb, Adida, and Moussa, "Reducing Misinformation"; Offer-Westort, Rosenzweig, and Athey, "Battling the Coronavirus Infodemic."

[31] Athey, et al., "Emotion-versus Reasoning-Based Drivers."

[32] Ali and Qazi, "Countering Misinformation."

[33] Badrinathan, "Educative Interventions"; Apuke, Omar, and Tunca, "Literacy Concepts," and Zhang, et al., "Experimenting the Effect."

[34] Guess, et al., "A Digital Media Literacy Intervention"; Offer-Westort, Rosenzweig, and Athey, "Battling the Coronavirus Infodemic."

[35] Athey, et al., "Emotion- versus Reasoning-Based Drivers."

[36] Apuke, Omar, and Tunca, "Literacy Concepts," and Zhang, et al., "Experimenting the Effect of Using Visual Multimedia Intervention."

[37] Ali and Qazi, "Countering Misinformation."

[38] Arechar, et al., "Understanding and Combatting Misinformation"; Apuke, Omar, and Tunca, "Literacy Concepts," and Zhang, et al., "Experimenting the Effect"; Offer-Westort, Rosenzweig, and Athey, "Battling the Coronavirus Infodemic."

[39] Badrinathan, "Educative Interventions"; Gottlieb, Adida, and Moussa, "Reducing Misinformation."

[40] Ali and Qazi, "Countering Misinformation" ; Guess, et al., "A Digital Media Literacy Intervention"; Athey, et al., "Emotion- versus Reasoning-Based Drivers."

[41] Athey, et al., "Emotion- versus Reasoning-Based Drivers."

[42] Ali and Qazi, "Countering Misinformation"; Guess, et al., "A Digital Media Literacy Intervention."

[43] Athey, et al., "Emotion- versus Reasoning-Based Drivers."

SUPPLEMENTAL INFORMATION

THE EVIDENCE SEARCH

To synthesize current evidence on the efficacy of media literacy interventions to address misinformation in the Global South, we conducted a literature search in November of 2023. The search was conducted on Google Scholar, Elicit, and ConnectedPapers to maximize the likelihood of including all relevant studies, both published and unpublished. The Google Scholar search terms were structured as follows, with backslashes indicating separate searches per term: "media literacy / digital literacy / information literacy / lateral reading / + misinformation / disinformation / malinformation / fake news / false news + Global South /developing country(ies) / East Asia / Latin America / Middle

East and North Africa / South Asia / sub-Saharan Africa." The questions feed to Elicit, an AI search engine which utilizes language models to pull relevant research from Semantic Scholar given a research question, which was as follows: "Are media, news, information, digital, and tech literacy interventions effective at countering misinformation?" The article given to Connected Papers (a visual tool which also uses Semantic Scholar to draw citation connections between papers) as the main reference node was Ford et al. (2023). Despite not using an RCT design, the paper was one of the most recently published on this topic in the Journal of Media Literacy Education and tests their intervention within a Global South country, making the study a useful starting point to identify other newly distributed studies on the same topic within the Global South

INDIVIDUAL STUDIES IN THE EVIDENCE BASE

Table 1: Summary of Global South Evidence

Study	Country	Participants	Interventions	Outcomes	Measured When	Findings(s)
Ali A., Qazi I.A. (2023)	Pakistan	Individuals in low and middle income households, recruited in person (n = 750)	Intermediate: Three minute in-person video with information about fake news and tips to identify it, could also include personalized enumerator feedback on real news stories that had been rated by respondents earlier	Belief discernment	Both immediate and longer term	*Mixed:* No effect for video only treatment; Video + feedback improved belief discernment among men and those with high digital literacy
Apuke O.D., Omar B., & Asude Tunca E. (2022) & Zhang, L., Iyendo T.O., Apuke O.D., & Gever C.D (2022)	Nigeria	Undergraduate students at one university (n = 470)	Intensive: Eight week training session covering information, news, and digital literacy	Both belief discernment and sharing intentions	Immediate	*Positive*: improved belief discernment and sharing intentions

Study	Country	Participants	Interventions	Outcomes	Measured When	Findings(s)
Arechar et al. (2023)	Argentina, Australia, Brazil, China, Egypt, India, Italy, Mexico, Nigeria, The Philippines, Russia, Saudi Arabia, Spain, United Kingdom, United States, South Africa	Individuals recruited online via social media, country-level quotas for age and sex (n = 34,286)	Light touch: Brief, one time exposure online to four simple digital literacy tips, taken from Facebook's tips	Belief discernment	Immediate	*Mixed: no effect for Global South countries; positive for some Global North countries and in the aggregate.*
Athey S., Cersosimo M., Koutout K., & Li Z. (Unpublished)	Kenya	English-speaking, adult Facebook users (n = 5,316)	Intermediate: One text per day for five days, containing a combination of emotional appeals, reason-based techniques, or a combination of emotion- and reason-based appeals	Both belief discernment and sharing intentions	Both immediate and longer term	*Positive: all three courses improve belief discernment and sharing intentions, with the emotions-based approach being the most effective*
Badrinathan S. (2021)	India	Individuals randomly selected from Gaya city in Bajar state (n = 1,224)	Intensive: An hour-long training on concrete tools to verify information plus personal demonstrations from research personnel	Belief discernment	Longer term	*Mixed: no effect on belief discernment in the aggregate; decreased belief discernment for supporters of national right-leaning party when stories are pro-attitudinal; improved belief discernment for non-right-wing supporters when stories are anti-attitudinal*
Gottlieb J., Adida C.L., & Moussa R. (Unpublished)	Côte d'Ivoire	18-30 years olds (n = 1,891)	Light touch: A single four minute video of a French journalist from Fact News presenting tips	Both belief discernment and sharing intentions	Longer term	*Null: no effect on discernment or sharing intentions*
Guess et al. (2020)	United States; India	India: Hindi-speaking individuals recruited online (n = 1,369) and in-person (n = 2,695)	Light touch: Brief, one time exposure to six strategies that readers can use to identify false or misleading stories, adapted from versions of messages taken from Facebook's tips and WhatsApp	Both belief discernment and sharing intentions	Both immediate and longer term	*Online intervention-Positive: improved discernment in immediate but not longer term outcomes* *In-person intervention-Null: no effect on belief discernment*
Offer-Westort, M., Rosenzweig, Leah R., & Athey, S. (Unpublished)	Kenya; Nigeria	Adults, recruited through Facebook ads (n = 10,531)	Light touch: Brief, one time exposure online to ten simple digital literacy tips, taken from Facebook's tips	Sharing intentions	Immediate	*Positive: improved sharing intentions*